

**PCT**WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau**B186**

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>6</sup> :</b> <b>C12N 15/12, C07K 14/47, A61K 38/17, G01N 33/68, 33/573, C12N 1/21</b>	<b>A2</b>	<b>(11) International Publication Number:</b> <b>WO 95/21252</b> <b>(43) International Publication Date:</b> 10 August 1995 (10.08.95)
<b>(21) International Application Number:</b> PCT/US95/01210 <b>(22) International Filing Date:</b> 31 January 1995 (31.01.95)  <b>(30) Priority Data:</b> 08/190,802      1 February 1994 (01.02.94)      US  <b>(71) Applicant:</b> THE BOARD OF TRUSTEES OF THE LELAND STANFORD JUNIOR UNIVERSITY [US/US]; Office of Technology Licensing, Suite 350, 900 Welch Road, Palo Alto, CA 94304 (US).  <b>(72) Inventors:</b> MOCHLY-ROSEN, Daria; 325 Harvard Avenue, Menlo Park, CA 94025 (US). RON, Dorit; 840 Alvarado Street, San Francisco, CA 94114 (US).  <b>(74) Agents:</b> MURASHIGE, Kate, H. et al.; Morrison & Foerster, 2000 Pennsylvania Avenue, N.W., Washington, DC 20006 (US).		<b>(81) Designated States:</b> AU, CA, JP, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).  <b>Published</b> <i>Without international search report and to be republished upon receipt of that report.</i>
<b>(54) Title:</b> WD-40-DERIVED PEPTIDES AND USES THEREOF		
<b>(57) Abstract</b>  The present invention relates to a polypeptide composition effective to alter the activity of a first protein that interacts with a second protein, where the second protein contains at least one WD-40 region. The polypeptides of the present invention typically have between 4 and 50 amino acids whose sequence is the same as a sequence of the same length in the WD-40 region of the second protein. The invention further includes a method of altering the activity of the above described first protein. In one embodiment of the invention the polypeptide composition is effective to alter the activity of a protein kinase C, where the protein kinase C interacts with a second protein, and the second protein contains at least one WD-40 region (e.g., RACK1).		

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

## METHODS AND COMPOSITIONS FOR MODULATION OF VESICULAR RELEASE

5           This work was supported in part by NIMH Grant 5R37MH38710 and NIH Grant 2P50MH48108. Accordingly, the United States Government has certain rights in this invention.

### Field of the Invention

10           The present invention relates to methods of identifying compounds capable of modulating vesicular release. In particular, the invention relates to (i) methods of identifying such compounds employing the secretion associated 17S (SA-17S) complex, and (ii) the proteins comprising the SA-17S complex.

### References

- 15           Aalto, M.K., *et al.*, *EMBO J.* 12:4095-4104 (1993).  
          Ausubel, F.M., *et al.*, *CURRENT PROTOCOLS IN MOLECULAR BIOLOGY*, John Wiley and Sons, Inc., Media, PA (1992).  
          Bennett, M.K., *et al.*, *J. Cell. Biol.* 116:761-775 (1992).  
          Bennett, M.K., *et al.*, *Cell* 74:863-873 (1993).  
20           Bowser, R., *et al.*, *J. Cell Biol.* 118:1041-1056 (1992).  
          Brennwald, P., and Novick, P., *Nature* 362:560-563 (1993).  
          Buckley, K., and Kelly, R.B., *J. Cell Biol.* 100:1284-1294 (1985).  
          Elferink, L.A., *et al.*, *Cell* 72:153 (1993).  
          Harlow, E., *et al.*, *ANTIBODIES: A LABORATORY MANUAL*, Cold Spring Harbor Laboratory  
25   Press (1988).  
          Lane, R.D., *et al.*, *Meth. Enzymol.* 121:183-192 (1986).  
          Malgaroli, A., and Tsien, R.W., *Nature* 357:134-139 (1992).  
          Mays, R.W., *et al.*, *J. Cell Biol.* 130:1105-1115 (1995).  
          Pevsner, J., *et al.*, *Neuron* 13:353-361 (1994a).  
30           Pevsner, J., *et al.*, *Proc. Natl. Acad. Sci. USA* 91:1445-1449 (1994b).  
          Scheller, R.H., *Neuron* 14:893-897 (1995).  
          Stone, K.L., and K.R. Williams, "Enzymatic Digestion of Proteins and HPLC Peptide Isolation" in *A PRACTICAL GUIDE TO PROTEIN AND PEPTIDE PURIFICATION FOR MICROSEQUENCING*,  
          2ND ED., Matsudaira, P., Ed., Academic Press, New York (1993).  
35           TerBush, D.R., and Novick, P., *J. Cell Biol.* 130:299-312 (1995).  
          TerBush, D.R., *et al.*, *EMBO J.* 15:6483-6494 (1996).  
          Ting, A.E., *et al.*, *Proc. Natl. Acad. Sci. USA* 92:9613-9617 (1995).

### Background of the Invention

Signal transmission between nerve cells typically involves the release of neurotransmitter from a presynaptic cell onto a postsynaptic cell. The neurotransmitter in the presynaptic cell is contained in synaptic vesicles positioned above the release sites at the presynaptic membrane (active zones). In response to a release signal (typically a local influx of calcium due to a depolarization of the presynaptic terminal), the vesicles undergo a series of mobilization steps culminating in the fusion of the vesicles with the presynaptic terminal membrane, and a dumping of vesicle contents into the synaptic cleft.

The neurotransmitter molecules diffuse across the synaptic cleft and bind to corresponding receptors in the postsynaptic membrane to communicate the appropriate signal (typically a depolarization or hyperpolarization of the postsynaptic membrane) to the postsynaptic cell. Much of the neurotransmitter in the synapse is subsequently re-absorbed by the presynaptic cell through specific transmitter uptake mechanisms.

A number of drugs affecting signalling in the central nervous system (CNS) and the peripheral nervous system (PNS) have been developed. Some of these, such as phenoxybenzamine, block specific post-synaptic receptors; others, such as clonidine and diethylamide, stimulate such receptors; still others (*e.g.*, desipramine, imipramine) act on reuptake mechanisms, and some act on neurotransmitter synthesis (*e.g.*,  $\alpha$ -Methyltyrosine, p-Chlorophenylalanine) or degradation (*e.g.*, monoamine oxidase inhibitors, iproniazid, pargyline).

The present invention provides a tool for the screening and identification of drugs capable of affecting secretory processes, such as neurotransmitter release at the active zones of presynaptic membranes.

### Summary of the Invention

In one aspect, the present invention includes a method of identifying a compound capable of modulating secretion of secretory vesicles. The method includes contacting an SA-17S complex with an SC complex, in the presence and absence of a test compound, measuring the effect of the test compound on the extent of binding between the SA-17S and the SC complexes, and identifying the compound as effective if its measured effect on the extent of binding is above a threshold level.

In one embodiment, the SA-17S complex binds to the 7S SC complex. In another embodiment, the SA-17S complex binds to the 20S SC complex.

The test compound may be effective to enhance (potentiate) or inhibit binding between the SA-17S and SC complexes. Compounds tested may include small molecules in a small molecule

combinatorial library, peptides (*e.g.*, peptides in a peptide combinatorial library), and the like. A compound which enhances binding between the SA-17S and SC complexes may be used to treat an affective disorder (such as depression, manic-depressive disorders and anxiety disorders) or a neurodegenerative disease (such as Parkinson's disease or Huntington's disease). A compound  
5 which inhibits binding between the SA-17S and SC complexes may be used to treat a disorder of thought, such as schizophrenia.

In another aspect, the invention includes a substantially purified SA-17S p71 polypeptide, having a molecular weight of about 71 kD and having the sequence of SEQ ID NO:54.

In another aspect, the invention includes a substantially purified SA-17S p79 polypeptide,  
10 having a molecular weight of about 79 kD and having the sequence of SEQ ID NO:56.

In another aspect, the invention includes a substantially purified SA-17S p84 polypeptide, having a molecular weight of about 84 kD and having the sequence of SEQ ID NO:58.

In another aspect, the invention includes a substantially purified SA-17S p96 polypeptide, having a molecular weight of about 96 kD and having the sequence of SEQ ID NO:60.

15 In another aspect, the invention includes a substantially purified SA-17S p102 polypeptide, having a molecular weight of about 102 kD and having the sequence of SEQ ID NO:62.

In another aspect, the invention includes a substantially purified SA-17S p106 polypeptide, having a molecular weight of about 106 kD and characterized by at least one sequence selected from the group consisting of SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID  
20 NO:49, SEQ ID NO:50, SEQ ID NO:51, and SEQ ID NO:52.

In another aspect, the invention includes a substantially purified polypeptide having (i) at least 60% sequence identity with a selected one of the SA-17S polypeptides p71, p79, p84, p96, p102, and p106, and (ii) the property of the selected polypeptide to form an SC-binding SA-17S complex when combined stoichiometrically with the other polypeptide components of the SA-17S  
25 complex.

In another aspect, the invention includes a substantially purified polynucleotide having the sequence SEQ ID NO:53 encoding the SA-17S p71 polypeptide.

In another aspect, the invention includes a substantially purified polynucleotide having the sequence SEQ ID NO:55 encoding the SA-17S p79 polypeptide.

30 In another aspect, the invention includes a substantially purified polynucleotide having the sequence SEQ ID NO:57 encoding the SA-17S p84 polypeptide.

In another aspect, the invention includes a substantially purified polynucleotide having the sequence SEQ ID NO:59 encoding the SA-17S p96 polypeptide.

In another aspect, the invention includes a substantially purified polynucleotide having the sequence SEQ ID NO:61 encoding the SA-17S p102 polypeptide.

In another aspect, the invention includes a substantially purified polynucleotide (i) having at least 60% sequence identity with a selected one of the polynucleotides encoding p71, p79, p84, p96,  
5 and p102, and (ii) encoding a polypeptide having the property of the polypeptide encoded by the selected polynucleotide to form an SC-binding SA-17S complex when combined stoichiometrically with other polypeptide components of the SA-17S complex.

In another aspect, the invention includes substantially purified antibodies, such as polyclonal or monoclonal antibodies, which are specifically-immunoreactive with one of the proteins or  
10 polypeptides described above.

These and other objects and features of the invention will become more fully apparent when the following detailed description is read in conjunction with the accompanying drawings.

#### Brief Description of the Figures

15 Fig. 1A shows a Coomassie blue-stained protein profile of glycerol gradient fractions from rat brain supernatant separated on a linear glycerol gradient, and analyzed by SDS-PAGE followed by Western blotting, with positions of rsec8, rsec6 and nsec1 on the gel indicated on the right as 8, 6, and 1 respectively. Fig. 1B shows immunodetection of rsec6, rsec8 and nsec1 in gradient fractions using immunoaffinity-purified antibodies, with size markers indicated by arrows. Fig. 1C  
20 shows co-immunoprecipitation of rsec6 and rsec8.

Figs. 2A, 2B and 2C show the purification of rat brain SA-17S complex. Fig. 2A shows Coomassie blue-stained SDS polyacrylamide gel analysis of pooled column fractions. Fig. 2B shows a Western blot analysis of the column fractions shown in Fig. 2A. Fig. 2C shows a Coomassie blue-stained SDS polyacrylamide gel analysis of purified SA-17S complex.

25 Figs. 3A, 3B, 3C and 3D show mass spectrometric analyzes of SA-17S complex peptides p102, p96, p71 and p84, respectively, with the mass of each peptide predicted by mass spectrometry shown above the peptide peak. The amino acid sequence of each peptide determined by peptide sequencing and its mass (protonated form of peptide) calculated from the determined peptide sequence is shown beside corresponding peptide peak.

30 Figs. 4A-4H show peptide sequences from the following SA-17S complex peptides, respectively: rsec8, p106, p102, p96, rsec6, p84, p79, and p71.

Fig. 5A shows a Western blot analysis of brain regional expression of the SA-17S complex. Fig. 5B shows the presence of the SA-17S complex in cultured cell lines, indicated.

Fig. 6A shows a Western blot analysis of subcellular fractionation of the SA-17S complex. Fig. 6B shows association of the SA-17S complex with brain membranes as analyzed by Western blot. Fig. 6C shows the nature of rsec8 association with brain membranes. Fig. 6D shows immunoprecipitation of synaptic vesicles from rat brain supernatant with anti-SV2 antibody or control mouse immunoglobulin coupled to Dynabeads M-500.

Figs. 7A-D shows cultures of primary hippocampal neurons fixed and immunostained with antibodies against rsec8 (Figs. 7A and 7B) and synaptotagmin (Fig 7 C). Fig. 7D shows a larger magnification of rsec8 staining overlaid with that of synaptotagmin at two synaptic terminals indicated by arrows. Bars = 7 $\mu$ m.

Fig. 8A is a Western blot analysis using seven monoclonal antibodies generated against rsec8 showing that syntaxin associates with the SA-17S complex. Fig. 8B shows a co-immunoprecipitation of syntaxin with rsec8.

#### Brief Description of the Sequences

- 15 SEQ ID NO:1 is the nucleotide sequence of rsec6.
- SEQ ID NO:2 is the amino acid sequence of rsec6.
- SEQ ID NO:3 is the nucleotide sequence of rsec8.
- SEQ ID NO:4 is the amino acid sequence of rsec8.
- SEQ ID NO:5 is the amino acid sequence of peptide fragment rsec6.1.
- 20 SEQ ID NO:6 is the amino acid sequence of peptide fragment rsec6.2.
- SEQ ID NO:7 is the amino acid sequence of peptide fragment rsec6.3.
- SEQ ID NO:8 is the amino acid sequence of peptide fragment rsec8.1.
- SEQ ID NO:9 is the amino acid sequence of peptide fragment rsec8.2.
- SEQ ID NO:10 is the amino acid sequence of peptide fragment rsec8.3.
- 25 SEQ ID NO:11 is the amino acid sequence of peptide fragment p71.1.
- SEQ ID NO:12 is the amino acid sequence of peptide fragment p71.2.
- SEQ ID NO:13 is the amino acid sequence of peptide fragment p71.3.
- SEQ ID NO:14 is the amino acid sequence of peptide fragment p79.1.
- SEQ ID NO:15 is the amino acid sequence of peptide fragment p79.2.
- 30 SEQ ID NO:16 is the amino acid sequence of peptide fragment p79.3.
- SEQ ID NO:17 is the amino acid sequence of peptide fragment p79.4.
- SEQ ID NO:18 is the amino acid sequence of peptide fragment p79.5.
- SEQ ID NO:19 is the amino acid sequence of peptide fragment p79.6.

- SEQ ID NO:20 is the amino acid sequence of peptide fragment p79.7.  
SEQ ID NO:21 is the amino acid sequence of peptide fragment p79.8.  
SEQ ID NO:22 is the amino acid sequence of peptide fragment p79.9.  
SEQ ID NO:23 is the amino acid sequence of peptide fragment p79.10.  
5 SEQ ID NO:24 is the amino acid sequence of peptide fragment p84.1.  
SEQ ID NO:25 is the amino acid sequence of peptide fragment p84.2.  
SEQ ID NO:26 is the amino acid sequence of peptide fragment p84.3.  
SEQ ID NO:27 is the amino acid sequence of peptide fragment p84.4.  
SEQ ID NO:28 is the amino acid sequence of peptide fragment p84.5.  
10 SEQ ID NO:29 is the amino acid sequence of peptide fragment p84.6.  
SEQ ID NO:30 is the amino acid sequence of peptide fragment p84.7.  
SEQ ID NO:31 is the amino acid sequence of peptide fragment p84.8.  
SEQ ID NO:32 is the amino acid sequence of peptide fragment p84.9.  
SEQ ID NO:33 is the amino acid sequence of peptide fragment p96.1.  
15 SEQ ID NO:34 is the amino acid sequence of peptide fragment p96.2.  
SEQ ID NO:35 is the amino acid sequence of peptide fragment p96.3.  
SEQ ID NO:36 is the amino acid sequence of peptide fragment p96.4.  
SEQ ID NO:37 is the amino acid sequence of peptide fragment p96.5.  
SEQ ID NO:38 is the amino acid sequence of peptide fragment p102.1.  
20 SEQ ID NO:39 is the amino acid sequence of peptide fragment p102.2.  
SEQ ID NO:40 is the amino acid sequence of peptide fragment p102.3.  
SEQ ID NO:41 is the amino acid sequence of peptide fragment p102.4.  
SEQ ID NO:42 is the amino acid sequence of peptide fragment p102.5.  
SEQ ID NO:43 is the amino acid sequence of peptide fragment p102.6.  
25 SEQ ID NO:44 is the amino acid sequence of peptide fragment p102.7.  
SEQ ID NO:45 is the amino acid sequence of peptide fragment p106.1.  
SEQ ID NO:46 is the amino acid sequence of peptide fragment p106.2.  
SEQ ID NO:47 is the amino acid sequence of peptide fragment p106.3.  
SEQ ID NO:48 is the amino acid sequence of peptide fragment p106.4.  
30 SEQ ID NO:49 is the amino acid sequence of peptide fragment p106.5.  
SEQ ID NO:50 is the amino acid sequence of peptide fragment p106.6.  
SEQ ID NO:51 is the amino acid sequence of peptide fragment p106.7.  
SEQ ID NO:52 is the amino acid sequence of peptide fragment p106.8.



- SEQ ID NO:53 is the nucleotide sequence encoding p71 (rsec10).  
SEQ ID NO:54 is the amino acid sequence of p71 (rsec10).  
SEQ ID NO:55 is the nucleotide sequence encoding p79 (rexo70).  
SEQ ID NO:56 is the amino acid sequence of p79 (rexo70).  
5 SEQ ID NO:57 is the nucleotide sequence encoding p84 (rexo84).  
SEQ ID NO:58 is the amino acid sequence of p84 (rexo84).  
SEQ ID NO:59 is the nucleotide sequence encoding p96 (rsec15).  
SEQ ID NO:60 is the amino acid sequence of p96 (rsec15).  
SEQ ID NO:61 is the nucleotide sequence encoding p102 (rsec5).  
10 SEQ ID NO:62 is the amino acid sequence of p102 (rsec5).

### Detailed Description of the Invention

#### I. Definitions

- "Substantially purified" refers to the at least partial purification of a selected polynucleotide,  
15 polypeptide, or related compound away from unrelated or contaminating components (*e.g.*, serum  
cells, other proteins).

- When a first polynucleotide fragment or polypeptide fragment is said to "correspond to" a  
second polynucleotide fragment or polypeptide fragment, respectively, it means that the fragments or  
regions are essentially co-extensive with one another when the sequences representing the fragments  
20 are aligned using a sequence alignment program, such as "MACVECTOR" (IBI, New Haven, CT).  
"Corresponding" polynucleotide or polypeptide fragments typically contain a similar, if not identical,  
number of residues. It will be understood, however, that corresponding fragments may contain  
insertions or deletions of residues with respect to one another, as well as some differences in their  
sequences.

- 25 The term "significant", when used with reference to "significantly different", "significantly  
inhibits" or "significantly stimulates", refers to a difference in a quantifiable parameter between the  
two groups being compared that is statistically-significant using standard statistical tests. For  
example, the degree of binding in a protein binding assay may be quantified using standard methods,  
and the degree of binding under different conditions can be compared for statistically-significant  
30 differences.

An antibody or antibody composition (*e.g.*, polyclonal antibodies) is "specifically  
immunoreactive" with a selected protein when the antibody or antibody composition is not reactive  
with antigens typically present in normal sera, not exposed to the selected protein.

A "syntaxin-containing complex", or "SC complex", refers to a complex formed of one or more proteins, where one of the proteins is syntaxin (Bennett, *et al.*, 1993; Scheller, 1995). Exemplary SC complexes include the 7S complex (comprising VAMP, synaptotagmin, syntaxin and SNAP-25) and the 20S complex (comprising VAMP, syntaxin, SNAP-25,  $\alpha$ -SNAP and NSF).

- 5 A polypeptide is "characterized by" a selected sequence when the polypeptide contains a sequence that is identical or substantially identical to the selected sequence.

"SNAP-25" refers to synaptosomal-associated protein of 25 kDa (Pevsner, *et al.*, 1994a; Scheller, 1995).

- 10  $\alpha$ -SNAP" refers to soluble NSF attachment protein (Pevsner, *et al.*, 1994a; Scheller, 1995).

"NSF" refers to *N*-ethylmaleimide-sensitive factor (Pevsner, *et al.*, 1994a; Scheller, 1995).

"VAMP" refers to vesicle-associated membrane protein (Pevsner, *et al.*, 1994a; Scheller, 1995).

15 II. Analysis, Isolation and Purification of the SA-17S Complex

- Experiments performed in support of the present invention and described in detail below have resulted in the identification of proteins that are involved in both regulated and constitutive exocytosis. Specifically, it has been discovered that in mammalian brain, the proteins rsec6 and rsec8, together with p71, p79, p84, p96, p102 and p106 proteins, are present in a high molecular weight complex, termed secretion-associated 17S (SA-17S) complex.

- 20 Analysis of the rat brain SA-17S complex revealed that it contains both rsec6 and rsec8. As described in Example 1, soluble brain proteins were fractionated on a linear glycerol gradient and the gradient fractions were analyzed by both sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and Western blotting. A Coomassie blue-stained protein profile of the fractions is shown in Figure 1A. The migration of rsec6 and rsec8 in the gradient was monitored by Western blot analysis using affinity-purified anti-rsec6 and anti-rsec8 antibodies. The results are shown in Figure 1B. Both rsec6 and rsec8 migrated at the 17S position. To determine if the co-fractionating rsec6 and rsec8 were in the same complex, rsec6 and rsec8 were immunoprecipitated from polyethylene glycol-fractionated rat brain supernatant with affinity-purified antibodies against rsec8 or with control rabbit immunoglobulin. The immunoprecipitated sample was analyzed for the presence of rsec6, rsec8 and nsec1 by Western blotting. The results, shown in Figure 1C, demonstrate that rsec6, but not nsec1, co-immunoprecipitated with rsec8 from soluble brain proteins using affinity-purified rabbit polyclonal anti-rsec8 antibodies. Taken together, these

data demonstrate that both rsec6 and rsec8 are components of the same high molecular weight complex.

Experiments detailed in Example 2 describe the purification of the SA-17S complex, and show that the complex is formed of 8 separate proteins or polypeptides. The SA-17S complex was purified from frozen rat brain by four sequential column chromatographic steps. The enrichment of the SA-17S complex following each chromatographic step was monitored by Western blot analysis using mouse polyclonal antibodies against rsec6 and rsec8, as shown in Figure 2B. After completing the steps detailed in Example 2, eight polypeptides of molecular weights 71, 79, 84, 86, 96, 102, 106, and 110 kDa were apparent in a Coomassie blue stained SDS-PAGE gel. These proteins were termed SA-17S p71, SA-17S p79, SA-17S p84, SA-17S p86, SA-17S p96, SA-17S p102, SA-17S p106, and SA-17S p110, respectively. Subsequent peptide sequence analyses, described below, determined that SA-17S p110 was rsec8 and SA-17S p86 was rsec6. SDS-PAGE analysis of pooled peak rsec6 and rsec8 immunoreactivity fractions indicated that these eight proteins constituted at least 95% of the total proteins in this fraction (Figure 2A, lane TMAE2; Figure 2C).

Fractionation of the purified complex on a glycerol gradient demonstrated that all eight components co-migrate at the 17S position. In addition, all eight proteins co-immunoprecipitated with rsec8 monoclonal antibodies. Further, all eight proteins were approximately equally stained by Coomassie blue, suggesting a stoichiometry of one copy of each protein within the complex (Figure 2C). This 1:1 stoichiometry is further supported by the apparent molecular weight of the native complex (~600-700 kD).

The SA-17S is also unusually stable, with monomeric forms of the proteins forming the complex never observed under non-denaturing conditions. Unlike the 20S particle composed of syntaxin, VAMP, SNAP-25,  $\alpha$ -SNAP and NSF, the SA-17S complex does not dissociate in response to ATP hydrolysis. In fact, this complex does not dissociate in the presence of ATP, GTP, ATP $\gamma$ S, GTP $\gamma$ S, EGTA and/or Ca<sup>2+</sup>. The mammalian SA-17S complex is found in soluble and membrane-bound states.

### III. Polypeptides Contained in the SA-17S Complex

To further characterize individual components of the mammalian SA-17S complex, the purified complex was fractionated on an SDS polyacrylamide gel and each protein band was cut out and individually subjected to in-gel proteolysis as described in Example 3, to generate peptides which could be further characterized by high performance liquid chromatography (HPLC), mass spectrometry for peptide purity and mass, and peptide sequencing.

Figures 3A, 3B, 3C and 3D show the mass spectrometric analyzes of HPLC peptide peaks from four proteins, SA-17S p102, SA-17S p96, SA-17S p84 and SA-17S p71, respectively, with the mass of each peptide predicted by mass spectrometry is shown above the peptide peak. The amino acid sequence of each peptide determined by peptide sequencing and its mass (protonated form of peptide) calculated from the determined peptide sequence is shown beside corresponding peptide peak, along with its SEQ ID NO. In Panels A to C the calculated and observed masses matched with less than 0.1% difference. Peptide sequences obtained from the eight members of the SA-17S complex are shown in Figures 4A-H, as well as Tables 1-8, below.

These sequences, as well as the molecular weights of the polypeptides, are effective to uniquely characterize the 8 proteins which form the SA-17S complex. In this respect, the invention includes the following substantially purified polypeptides: (i) SA-17S p71 polypeptide, having a molecular weight of about 71 kD and characterized by at least one sequence selected from the group consisting of SEQ ID NO:11, SEQ ID NO:12 and SEQ ID NO:13; (ii) SA-17S p79 polypeptide, having a molecular weight of about 79 kD and characterized by at least one sequence selected from the group consisting of SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22 and SEQ ID NO:23; (iii) SA-17S p84 polypeptide, having a molecular weight of about 84 kD and characterized by at least one sequence selected from the group consisting of SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, and SEQ ID NO:32; (iv) SA-17S p96 polypeptide, having a molecular weight of about 96 kD and characterized by at least one sequence selected from the group consisting of SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36 and SEQ ID NO:37; (v) SA-17S p102 polypeptide, having a molecular weight of about 102 kD and characterized by at least one sequence selected from the group consisting of SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:43, and SEQ ID NO:44; and (v) SA-17S p106 polypeptide, having a molecular weight of about 106 kD and characterized by at least one sequence selected from the group consisting of SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, and SEQ ID NO:52.

These peptides are useful for the formation of a complete SA-17S complex, which can be employed, *e.g.*, in screens such as the screen described below. The peptides may also be used to identify the site of interaction between the SA-17S and SC complexes, enabling the design of a simpler screen for compounds which modulate such an interaction.

A. Isolation of Full-length Clones Encoding SA-17S Polypeptides

The peptide sequences provided in Tables 1-8, below, were used to isolate DNA or cDNA clones encoding the complete proteins, according to standard techniques. From these studies, polynucleotide sequences which encode full-length SA-17S polypeptides p71, p79, p84, p96, and p102. Based on amino acid sequence similarities, six of the eight components of the mammalian SA-17S complex have counterparts in the 834 kDa yeast Sec6/8/15 complex (TerBush and Novick, 1995; TerBush, *et al.*, 1996).

The SA-17S p71 cDNA (SEQ ID NO:53) has an open reading frame (ORF) which encodes a 708 amino acid polypeptide (SEQ ID NO:54) with a calculated molecular weight of 82 kdal. The p71 polypeptide is homologous to yeast sec10 and is thus denoted rsec10. The amino acid sequence contains potential coiled-coil domains at the N-terminus and near the C-terminus of the protein. In addition, rsec10 is homologous with the tail domain of yeast type II myosin, and may interact with cytoskeletal proteins.

The SA-17S p79 cDNA (SEQ ID NO:55) has an ORF which encodes a 653 amino acid protein (SEQ ID NO:56) with a calculated molecular weight of 75 kDa. The p79 polypeptide sequence shows 23% identity and 33% similarity to yeast exo70 and is thus denoted rexo70. Using the Coils 2.1 program, a region of 30 amino acids (residues 5-34) at the amino terminus of the protein is predicted to have a high probability (between 0.88 and 1.0) of forming a coiled-coil domain, a structure which is known to be involved in protein-protein interactions.

The SA-17S p96 cDNA (SEQ ID NO:59) has an ORF which encodes an 822 amino acid protein (SEQ ID NO:60) with a calculated molecular weight of 95 kDa. The p96 protein sequence shows 24% identity and 36% similarity to yeast sec15 and is designated rsec15. A region of 22 amino acids (residues 85-106) at the amino terminus of the protein has a probability of greater than 0.59 to form a coiled-coil domain.

The SA-17S p102 cDNA (SEQ ID NO:61) has an ORF which encodes a 924 amino acid protein (SEQ ID NO:62) with a calculated molecular weight of 104 kDa. The p102 protein sequence shows 21% identity and 32% similarity to yeast sec5, and is denoted rsec5. A region of 21 amino acids (residues 240-260) at the amino terminus of the protein has a high probability (greater than 0.94) to form a coiled-coil domain.

The cDNA sequence of the SA-17S p84 subunit (SEQ ID NO:57) was identified based on overlapping rat cDNA clones which encode the peptide sequences derived from the p84 subunit (SEQ ID NO:24-32). The predicted protein (SEQ ID NO:58) is designated rexo84 since it corresponds to the 84 kDa protein in the complex. Like other proteins of the SA-17S complex,

rexo84 contains regions predicted to form coiled-coils. The rexo84 sequence is not significantly similar to any of the known components of the 834 kDa yeast Sec6/8/15 complex.

All the above-identified members of the mammalian SA-17S complex, as well as rsec6 and rsec8, contain regions predicted to form coiled-coil domains, suggesting that subunits in this  
5 complex interact with each other at least in part via these domains. Predicted coiled-coil domains are prominent in other membrane trafficking proteins including the SNAREs, suggesting these interactions may be involved in inter-complex interactions as well.

#### B. Recombinant Production of the SA-17S Proteins

10 Polynucleotide sequences encoding proteins of the SA-17S complex may be cloned into an expression plasmid, such as a p-GEX plasmid (Pharmacia), to produce corresponding polypeptides. Recombinant pGEX plasmids can be transformed into appropriate strains of *E. coli* and fusion protein production can be induced by the addition of IPTG (isopropyl-thio galactopyranoside). Solubilized recombinant fusion protein can then be purified from cell lysates of the induced cultures  
15 using glutathione agarose affinity chromatography according to standard methods (Ausubel, *et al.*, 1992).

Affinity chromatography may also be employed for isolating  $\beta$ -galactosidase fusion proteins (such as those produced by lambda gt11 clones). The fused protein is isolated by passing cell lysis material over a solid support having surface-bound anti- $\beta$ -galactosidase antibody.

20 Isolated recombinant polypeptides produced as described above may be purified by standard protein purification procedures. These procedures may include differential precipitation, molecular sieve chromatography, ion-exchange chromatography, isoelectric focusing, gel electrophoresis and affinity chromatography.

In addition to recombinant methods, SA-17S complex proteins or polypeptides can be  
25 isolated from selected cells by affinity-based methods, such as by using appropriate antibodies (described below). Further, SA-17S complex peptides may be chemically synthesized using methods known to those skilled in the art.

#### C. Antibodies Directed Against SA-17S Polypeptides

30 SA-17S polypeptides of the present invention, such as polypeptides identified by sequences SEQ ID NO:5-52, as well as full-length polypeptides corresponding to one of the eight protein components of SA-17S, such as polypeptides identified by sequences SEQ ID NO:2, 4, 54, 56, 58, 60, and 62, and fragments thereof, may be used in the generation of antibodies, *e.g.*, as described in

the Materials and Methods below. The polypeptides may be used in unmodified form, or they may be coupled to appropriate carrier molecules, such as bovine serum albumin (BSA) or Keyhole Limpet Hemocyanin (KLH) (available from, for example, Pierce, Rockford, IL).

To prepare antibodies, a host animal, such as a mouse or rabbit, is typically immunized with the purified polypeptide, or polypeptide coupled to carrier or fusion protein (*e.g.*, a His-tagged fusion protein or a glutathione-S-transferase fusion). The host serum or plasma is collected following an appropriate time interval, and the serum is tested for antibodies specific against the polypeptide.

The gamma globulin fraction or the IgG antibodies of immunized animals can be obtained, for example, by use of saturated ammonium sulfate precipitation or DEAE Sephadex chromatography, affinity chromatography, or other techniques known to those skilled in the art for producing polyclonal antibodies.

Alternatively, purified antigenic polypeptide or fused antigen protein may be used for producing monoclonal antibodies. In this case, the spleen or lymphocytes from an immunized animal are removed and immortalized or used to prepare hybridomas by methods known to those skilled in the art (*e.g.*, Harlow, *et al.*, 1988). For example, hybridomas may be prepared by fusion of NS-1 mouse myeloma cells with spleen cells from BALB/c mice immunized with His-tagged fusion protein (Lane, *et al.*, 1986). Antibodies secreted by the immortalized cells are screened (*e.g.*, using enzyme linked immunosorbent assay (ELISA) or a Western blot) to determine the clones that secrete antibodies of the desired specificity (*e.g.*, Ausubel, *et al.*, 1992).

If desired, the antibodies may be affinity-purified prior to use, using methods known in the art (*e.g.*, Harlow, *et al.*, 1988). Antibodies or FAb fragments thereof generated as described above may be used in a variety of ways. In particular, they can be used to detect or quantitate the level of SA-17S polypeptides in any of the methods of the present invention where such detection is contemplated, *e.g.*, in methods employing Western-based detection approaches. Further, the antibodies can be used in screens of expression libraries, as described above. The antibodies may also be used to co-immunoprecipitate proteins which interact with the SA-17S complex, as described in the Materials and Methods and Example 7.

#### IV. Expression Pattern of the SA-17S Complex

Expression patterns of the SA-17S complex are described below in Examples 4-6. Peripheral rsec8 staining in synaptic terminals suggests that the SA-17S complex is associated with the synaptic plasma membrane. The presence of the SA-17S complex in hippocampal synapses,

together with its association with the synaptosomal protein syntaxin, indicate that the SA-17S complex plays a critical role in the synaptic vesicle trafficking pathway of yeast and mammals.

#### V. Role of the SA-17S Complex in Secretion

5       The SA-17S complex is required for normal secretion and synaptic release, and therefore, compounds which inhibit the association of the SA-17S complex with syntaxin or a syntaxin-containing complex will inhibit the release of the secretory vesicles.

      The results of the experiments described herein, including (i) the demonstration of an *in vivo* interaction between rsec8 and syntaxin as revealed by the immunoprecipitation of syntaxin with  
10 anti-rsec8 antibodies, and (ii) selective staining of neurons with anti-rsec8 antibodies in the area near the plasma membrane of the nerve terminals (the site of exocytosis), are in agreement with studies of yeast secretory mutants. The yeast studies have revealed genetic interactions between Sec8/Sec15 and the yeast homologs of syntaxin and SNAP25, SSO1 and Sec9. Specifically, overexpression of SSO1 can suppress mutations in the SEC15 gene (Aalto, *et al.*, 1993), while over-expression of Sec9  
15 can suppress mutations in both the SEC8 and SEC15 genes (Brennwald and Novick, 1993). Additionally, staining of the Sec6/8/15 complex in yeast appears enhanced at the site of vesicle docking and fusion in the tip of the bud.

      The data presented here are also consistent with a role for the SA-17S complex in a Rab mediated event at the plasma membrane, possibly intervening between the Rab and the 7S complex  
20 (Bowser, *et al.*, 1992).

      Expression of the SA-17S complex is not as high in neurons and endocrine cells as some of the isoforms of vesicle trafficking proteins involved in the synaptic vesicle docking and fusion pathway. Northern and Western blot analyses as well as fluorescence microscopy studies detect its expression in all tissues and cultured cell lines examined. This ubiquitous distribution of the SA-17S  
25 complex suggests that it is important in both constitutive and regulated membrane trafficking and that its function is fundamental to the exocytotic process.

      Although the SA-17S complex is described herein as pertaining primarily to neurotransmitter release, it is contemplated that the SA-17S complex is involved in secretory release in a number of other cell types, including mast cells.

30

#### VI. Screens for Inhibitors of Synaptic Protein Interactions

      The present invention includes methods of screening for compounds effective to modulate vesicular release involved in synaptic transmission and other secretory processes. The methods are



used to identify a compound capable of affecting binding of a secretion associated 17S (SA-17S) complex to a syntaxin-containing (SC) complex. An SA-17S complex is contacted with an SC complex, in the presence and absence of a test compound. The effect of the test compound on the extent of binding between the SA-17S and the SC complexes is measured, and a compound is  
5 identified as effective if its effect on the extent of binding is above a threshold level (*e.g.*, a several-fold difference in binding level between control and experimental samples).

The test compound may be effective to enhance (potentiate) or inhibit binding between the SA-17S and SC complexes. Compounds tested may include small molecules in a small molecule combinatorial library, peptides in a peptide combinatorial library, and the like.

10 Compounds which affect the binding of the SA-17S and SC complexes to one another, when applied to target cells, are expected to modulate vesicular release by those cells. The modulation may be an inhibition of release or stimulation of release, either when the compound is applied alone, or when the compound is applied in conjunction with another compound having an effect on vesicular release.

15 In one embodiment, the assay is conducted using a co-immunoprecipitation method such as is described in the Materials and Methods and Example 7. Anti-rsec8 monoclonal antibodies (such as antibodies 2E9 and/or 17A10) or mouse immunoglobulin are coupled to protein G beads at a final concentration of 2 mg/ml using the cross-linker dimethylpimelimidate (Pevsner, *et al.*, 1994a). Rat brain membranes prepared as described herein are resuspended in 20 mM Tris, pH 8.0, 150 mM  
20 NaCl and solubilized with either 1% CHAPS or Triton X-100 at a final protein concentration of 2 mg/ml.

The solubilized sample is incubated at 4°C for 30 min before centrifuging at 20,000 xg for 20 min. Following centrifugation, the solubilized brain membranes are pre-cleared by incubation with protein G beads (1 ml solubilized membranes per 200  $\mu$ l protein G beads) for 30 min at 4°C.  
25 The pre-cleared supernatant is then incubated with (i) a selected concentration of the test compound and (ii) 20  $\mu$ l protein G beads with either immobilized anti-rsec8 monoclonal antibodies 2E9/17A10 or immobilized control mouse Ig for 4 hrs at 4°C. The beads are then washed three times with 300  $\mu$ l 20 mM Tris, pH 8.0, 150 mM NaCl containing either 0.7% CHAPS or Triton X-100. Proteins bound to the beads are analyzed by Western blotting for the presence of syntaxin as described in the  
30 Materials and Methods and Examples below.

If the test compound is effective to inhibit binding of an SC complex to the immobilized SA-17S complex, the amount of syntaxin detected in the Western blot will be diminished. Conversely, if the test compound is effective to enhance the binding of an SC complex to the immobilized SA-

17S complex, the amount of syntaxin detected in the Western blot will be increased. The assay is typically conducted with both negative controls (*e.g.*, beads coated with control mouse Ig) and positive controls (no test compound in assays with anti-rsec8 antibody).

It will be appreciated that the co-immunoprecipitation assay may be practiced using other  
5 antibodies which recognize and can immunoprecipitate the SA-17S complex, such as anti-rsec6 antibodies or antibodies directed against one or more of the other proteins in the complex.

A. Measuring the Effect of a Test Compound on the Extent of Binding

The type of measurement used to quantify the effect of a test compound on the extent of  
10 binding between the SA-17S complex and a SC complex depends on the type of screening assay and detection system used, and can be readily determined by one of skill in the art.

For example, in a co-immunoprecipitation screen employing a Western blot detection, such as described above, the extent of binding may be measured using densitometry of the Western blot image. The densitometry values are typically normalized, and a threshold level is set based on the  
15 amount of variation in the signal between a series of "control" samples (samples not containing test compounds). The smaller the variation, the smaller the effect of a test compound that can be reliably detected. The threshold is typically set at a several-fold difference, such as a 3-5 fold increase or decrease in binding affinity.

The precise threshold level used in a particular application of the invention is determined in  
20 view of the specific requirements of that particular application. For example, if it is desired to isolate only compounds with a very high activity, the threshold is set to a relatively high value, such as a 10 to 100-fold difference in binding affinity. If, on the other hand, it is desired to isolate compounds having a subtle effect on the binding, a lower threshold level may employed.

It is also appreciated that, as with any other compound screening assay, the effect of a  
25 particular compound on the binding of a SA-17S complex to an SC complex depends on the concentration of the compound. At relatively high compound concentrations the effect may be large, and be manifested as, *e.g.*, a 50-fold difference in binding between control and experimental samples, whereas at lower compound concentrations, the effect may be smaller.

30 B. Effects of Identified Compounds on Vesicular Release

Compounds identified by a screen such as one described above as affecting the binding of an SA-17S complex to an SC complex may be further evaluated for their ability to modulate vesicular release *in vitro* and *in vivo*.

For example, the compounds may be tested using the PC12 cell D $\beta$ H vesicular release assay (Bennett, *et al.*, 1993), which detects a membrane-associated form of the enzyme dopamine  $\beta$ -hydroxylase (D $\beta$ H) on the luminal side of catecholamine-containing granules. When the cells are depolarized in the presence of calcium, granule fusion with the plasma membrane results in the exposure of D $\beta$ H on the cell surface, where it can be quantitatively detected by immunofluorescence microscopy (Elferink, *et al.*, 1993). By treating a sample of cells with a compound identified as affecting the binding of an SA-17S complex to an SC complex, depolarizing the cells (*e.g.*, with a pulse of KCl) in the presence of calcium, and comparing the response to that obtained with an untreated sample of cells, the effects of the compound on vesicle release in PC12 cells may be assessed. Similar assays may be employed using freshly-isolated cells (*e.g.*, in brain slices), or suitable animal models.

#### VII. Compounds Suitable for Screening

A variety of different compounds may be screened using methods of the present invention. They include peptides, macromolecules, small molecules, chemical and/or biological mixtures, and fungal, bacterial, or algal extracts. Such compounds, or molecules, may be either biological, synthetic organic, or even inorganic compounds, and may be obtained from a number of sources, including pharmaceutical companies and specialty suppliers of libraries (*e.g.*, combinatorial libraries) of compounds.

A set of potentially-effective test peptides can be generated from overlapping peptides spanning the entire sequence of each of the proteins involved in the SA-17S complex/SC complex interaction. Such a set is likely to contain peptides which may be effective to disrupt the interactions of the SA-17S complex with the SC complex.

In cases where an identified active compound is a peptide, the peptide may be utilized to aid in the discovery of orally-active small molecule mimetics.

#### VIII. Applications

Inhibitory compounds isolated using methods of the present invention may be employed to inhibit vesicle-mediated secretion of molecules from cells. Similarly, compounds which enhance or potentiate the binding of SA-17S complex to syntaxin or an SC complex may be used to upregulate vesicle-mediated secretion. The ability to modulate secretion processes has utility in a variety of areas, some of which are identified below.

A. CNS Disease Applications

A number of disorders and/or conditions of the central nervous system (CNS) may be alleviated by selectively enhancing or inhibiting vesicular release in specific areas of the brain. They include affective disorders (*e.g.*, depression), disorders of thought (*e.g.*, schizophrenia) and  
5 degenerative disorders (*e.g.*, Parkinson's disease), as well as applications such as anesthesia. A variety of drugs are currently used to treat such disorders and/or conditions. Compounds identified by methods of the present invention may be used either alone, or in combination with currently used therapies to alleviate symptoms associated with the disorders.

Drugs used to treat affective disorders, which include depression, manic-depressive disorders  
10 and anxiety disorders, typically fall into three classes: (i) monoamine oxidase (MAO) inhibitors, such as phenelzine, (ii) tricyclic compounds, such as imipramine, and (iii) serotonin uptake blockers, such as fluoxetine and trazodone. All of these drugs work, at least in part, by increasing the concentration of either serotonin or biogenic amine neurotransmitters in CNS synapses of treated individuals. According to methods of the present invention, compounds which enhance or stimulate  
15 the release of serotonin or biogenic amines at selected brain synapses may be similarly effective at treating affective disorders. Such compounds may be identified by screening for compounds which enhance the binding of SA-17S or its components with an SC complex.

Disorders of thought, such as schizophrenia, have been treated with a variety of antipsychotic drugs (including phenothiazines, such as chlorpromazine, butyrphenones, such as  
20 haloperidol, xithioxanthenes, and newer drugs, such as clozapine) now known to act as blockers of dopamine receptors. According to the teachings presented herein, compounds identified as inhibitors of release of dopamine-containing vesicles, particularly vesicles released from cells having their cell bodies in the arcuate nucleus of the hypothalamus, the substantia nigra, or the ventral tegmental area, may be employed to relieve symptoms of schizophrenia. Such compounds may be identified  
25 using a screening assay such as described above, for compounds effective to inhibit the binding of SA-17S complex or its components with an SC complex.

Neurodegenerative diseases, such as Parkinson's disease and Huntington's disease, may also benefit from compounds identified according to the methods of the present invention. Parkinson's disease arises due to degeneration of the nigrostriatal pathway, raphaei nuclei, locus ceruleus, and  
30 motor nucleus of vagus, which result in a reduction of dopamine, serotonin and norepinephrine levels. The symptoms of Parkinson's may be alleviated by administering compounds identified according to the teachings presented herein as stimulating release of vesicles containing the above neurotransmitters.

### B. Other Applications

In addition to applications in the CNS, compounds identified employing methods of the present invention may be used to therapeutically intervene in a variety of other systems. They include the endocrine system for treatment of hormonal imbalances, the immune system for intervention in antigen processing, secreted immunomodulators, and viral processing, as well as anti-tumor applications, such as regulation of membrane trafficking during rapid cell division.

The following examples illustrate but in no way are intended to limit the present invention.

### Materials and Methods

Unless otherwise indicated, [<sup>125</sup>I]goat anti-rabbit secondary antisera and the Enhanced Chemo-Luminescence (ECL) system were obtained from Amersham Corp. (Arlington, Heights, IL). Nitrocellulose paper was obtained from Schleicher and Schuell (Keene, NH). Materials for SDS-polyacrylamide gel electrophoresis (SDS-PAGE) were obtained from Bio-Rad Laboratories (Hercules, CA). Lys C and Glu C were obtained from Waco. Other chemicals were purchased from Sigma (St. Louis, MO) or United States Biochemical (Cleveland, OH).

Anti-rsec8 monoclonal antibodies (mAbs) may be generated as described herein, or may be purchased from commercial sources (*e.g.*, Transduction Laboratories, Lexington, KY).

All protein and protein complex purification procedures were carried out at 4°C unless otherwise noted.

### A. Glycerol Gradient Analysis

Rat brain supernatant was fractionated on a linear glycerol gradient (Ting, *et al.*, 1995) and analyzed by Western blotting. Rat brain supernatant was prepared by homogenization in 6 ml of 40 mM Hepes (pH 7.4), 150 mM NaCl, 10% (vol/vol) glycerol, 1 mM EDTA, 1 mM dithiothreitol containing leupeptin, aprotinin, pepstatin, and phenylmethanesulfonyl fluoride each at 1 µg/ml. The homogenate was centrifuged at 1000 ×g for 5 min and the supernatant was centrifuged at 100,000 ×g for 20 min. Eighty microliters (40 µg of protein) of the resulting supernatant was layered onto a 1.2-ml ten-step 22.5-36% (vol/vol) glycerol gradient in 40 mM Hepes (pH 7.4), 150 mM NaCl, 1 mM dithiothreitol. The gradient was centrifuged at 91,000 ×g for 16 hours at 4°C.

Fractions (95 µl) were collected and analyzed by SDS/PAGE and Western blotting. Migration of brain rsec8, rsec6 and n-Sec1 was monitored by Western blotting using affinity-purified antibodies.

B. Purification of Rat Brain SA-17S Complex

1. Bio-Gel HT Hydroxyapatite Chromatography. Fifty frozen rat brains (Harlan, IN) were homogenized in 400 ml of homogenization buffer (20 mM Hepes, pH 7.4, 200 mM NaCl, 1 mM dithiothreitol (DTT), 0.1 mM EDTA and 0.3 mM PMSF) by eight up-and-down strokes with a Teflon/glass homogenizer. A postnuclear supernatant was obtained by centrifuging the homogenate at 5000 rpm in a JA-10 rotor (Beckman, CA) for 15 min. Soluble brain protein supernatant was obtained by centrifuging the postnuclear supernatant at 36,000 rpm in a Ti45 rotor (Beckman, CA) for one hour.

Following the centrifugation, a final concentration of 110 mM sodium phosphate, pH 7.4 was added to the supernatant before it was applied to a 25 ml hydroxyapatite (BioRad, CA) column equilibrated with the homogenization buffer containing 110 mM sodium phosphate at a flow rate of 20 ml/hr at 18°C. The column was washed with 75 ml of 0.15 M sodium phosphate in buffer A (0.15 M NaCl, and 1 mM DTT, pH 7.4) and eluted stepwise with 5 ml aliquots of 0.2M to 0.65 M phosphate in buffer A, with 50 mM phosphate increment per step. Fractions containing rsec6 and rsec8 as determined by Western blot analyzes were pooled and dialyzed twice against 2 liters of 20 mM Tris, pH 8.0, 100 mM NaCl, 0.2 mM EDTA, and 0.5 mM DTT overnight.

2. Fractogel EMD TMAE-650 (S) Chromatography. The dialyzed hydroxyapatite eluant was diluted with an equal volume of 20 mM Tris, pH 8.0 and loaded onto a 3 ml column of Fractogel TMAE anion exchange resin (EM Separations Technology, NJ) equilibrated with 20 mM Tris, pH 8.0, 50 mM NaCl, and 1 mM DTT at a flow rate of 15 ml/hr. The column was washed with 20 mM Tris, pH 8.0, 50 mM NaCl, 1 mM DTT and eluted with a linear gradient of 0 to 350 mM NaCl in 20 mM Tris, pH 8.0 and 1 mM DTT.

3. HW-55S gel Filtration Chromatography. TMAE eluant fractions enriched in rsec6 and rsec8 were pooled and fractionated over a 180 ml HW-55S gel filtration column (TosoHaas, PA) equilibrated in 20 mM Hepes, pH 7.4, 200 mM NaCl, and 1 mM DTT at a flow rate of 5 ml/hr.

4. Second Fractogel EMD TMAE-650 (S) Chromatography. rsec6 and rsec8-enriched fractions from HW-55S gel filtration column chromatography were pooled and diluted to a final buffer concentration of 20 mM Tris, pH 8.0 and 50 mM NaCl. The diluted sample was applied to and eluted from a 0.5 ml Fractogel TMAE-650 column as described above.

Fractions containing purified SA-17S complex were pooled and used for SDS-PAGE analyzes and peptide sequencing.

C. Mass Spectrometry and Peptide Sequence Analysis

- 5 Purified SA-17S complex was fractionated on an 8% SDS polyacrylamide gel. Individual protein bands were cut out and subjected to in-gel proteolysis by Lys C, trypsin and/or Glu C essentially as described by Stone and Williams (1993). Excised gel slices were dried by Speed-Vac, and 500  $\mu$ l of 0.1 M  $\text{NH}_4\text{HCO}_3$  was added to each gel slice. The slices were shaken at room temperature for at least 4 hours. The wash from each gel slice was removed and saved. An additional 500  $\mu$ l of 0.1 M  $\text{NH}_4\text{HCO}_3$  was added to each gel slice, incubated overnight with shaking, and the wash removed and saved. A final 150  $\mu$ l of 0.1 M  $\text{NH}_4\text{HCO}_3$  was added to each gel slice, plus 5  $\mu$ l of 45 mM dithiothreitol. The gel slices were incubated at 50°C for 20 min., and cooled to room temperature. After addition of 5  $\mu$ l 100 mM iodoacetamide, the slices were incubated in the dark at room temperature for 20 min. Proteolytic enzyme (Lys C, trypsin and/or Glu C) was added in a 1:25 weight-to-weight ratio of enzyme:protein (with a typical gel slice containing 8-10  $\mu$ g of protein). After incubating the gel slices overnight at 37°C, 500  $\mu$ l of 0.1 M  $\text{NH}_4\text{HCO}_3$  was added. The gel slices were shaken for 8 hr to extract the peptides. The supernatant was transferred to a tube, and 300  $\mu$ l 2M urea in 0.1 M  $\text{NH}_4\text{HCO}_3$  was added to each gel slice. After 24 hr, the supernatant was transferred to the corresponding tube. The combined supernatants were dried in a Speed-Vac and redissolved in 210  $\mu$ l  $\text{H}_2\text{O}$ . The digested peptides were fractionated by HPLC (Applied Biosystems). Selected peptide peaks were then subjected to amino acid sequencing by Edman degradation method (Applied Biosystems, CA). Following digestion, some peptides from p106, p96 and p79 were also fractionated on a 1090 HPLC (Hewlett-Packard). Selected HPLC fractions were subjected to MALDI-TOF mass analysis (Hewlett-Packard). Following mass spectrometric analysis, selected peptide fractions were subjected to amino acid sequencing (Hewlett Packard).

D. Brain Regional and Tissue Culture Cell Western Blot Analysis

- For brain regional Western blot analysis, various rat brain regions were dissected from fresh brain and sonicated in 10 mM Hepes, pH 7.4, 2.5 mM KOAc, 1 mM  $\text{MgCl}_2$ , 0.1 mM EGTA, and 0.3 mM PMSF. The homogenates were centrifuged at 10,000 xg for 5 min and the postnuclear supernatants were collected.

For Western blot analysis of the rsec8 distribution in various cultured cells, tissue culture cells were grown to confluency in 10 cm petri dishes, scraped, and dounce homogenized in 20 mM Hepes, pH 7.4. The homogenates were centrifuged at 5000 xg for 5 min and the supernatants were collected for Western blot analysis.

- 5 Protein samples were electrophoresed on 12.5% resolving SDS-polyacrylamide (denaturing) gels, transferred to nitrocellulose paper (0.2  $\mu$ M), and probed as described below. Proteins were visualized by enhanced chemiluminescence (ECL) system (Amersham Corp., Arlington, Heights, IL) and/or by autoradiography using  $^{125}$ I-labeled goat anti-rabbit antibodies as the secondary antibody. Protein bands were quantitated by densitometry and/or phosphorimaging (Molecular Dynamics,  
10 Sunnyvale, CA).

#### E. Subcellular Localization of SA-17S Complex

- For brain subcellular fractionation studies, postnuclear supernatant and crude synaptosomal membrane fractions were prepared as described (Bennett, *et al.*, 1992). Soluble brain proteins and  
15 brain membranes were obtained by centrifuging the postnuclear supernatant or lysed crude synaptosomal fraction at 100,000 xg for 20 min in a TLS 55 rotor (Beckman, CA) at 4°C. All protein samples were subjected to SDS-PAGE and Western blot analysis.

- For brain membrane extraction studies, lysed brain membranes resuspended in 20 mM Tris, pH 8.0, 1 mM DTT and 10% sorbitol (2 mg/ml) were incubated with equal volumes of either 20  
20 mM Tris, pH 8.0, 3 M NaCl in 20 mM Tris, pH 8.0, 8% Triton X-100 in 20 mM Tris, pH 8.0, or 8 M urea in 20 mM Tris, pH 8.0. The incubations were carried out at 4°C for 1 hr. Following the incubation, the membranes were pelleted at 50,000 xg. The resulting supernatants and membrane pellets were analyzed for their rsec8 content by Western blot.

- The amount of rsec8 extracted from membranes were quantitated by phosphorimaging using  
25  $^{125}$ I-labeled secondary antibodies. The pixel values for 20 mM Tris extraction were 6713 for supernatant and 173171 for pellet; for 1.5 M NaCl extraction were 39189 for supernatant and 45414 for pellet; for Triton X-100 extraction were 56302 for supernatant and 37254 for pellet; and for urea extraction were 71419 for supernatant and 9999 for pellet.

- For sucrose gradient flotation analysis of SA-17S complex, lysed crude synaptosomal  
30 membranes were resuspended in 55% sucrose in gradient buffer (20 mM Hepes, pH 7.4, 150 mM NaCl and 1 mM DTT) at a protein concentration of 4 mg/ml. The resuspended membranes were placed at the bottom of a linear 25-52.5% sucrose gradient. After centrifugation at 42,000 rpm for 16 hrs in a TLS 55 rotor (Beckman, CA), the gradient was fractionated and analyzed by Western



blot for rsec6, rsec8, Na/K ATPase (a plasma membrane marker) and synaptophysin (a synaptic vesicle marker).

F. Immunoprecipitation Studies

5                   1. Immunoprecipitation of the SA-17S Complex. Immunoaffinity purified anti-rsec8 antibodies (Ting, *et al.*, 1995) or purified rabbit immunoglobulin were coupled to protein A beads at a final concentration of 0.5 mg/ml as previously described (Pevsner, *et al.*, 1994a; Pevsner, *et al.*, 1994b). Three ml of rat brain supernatant (Ting, *et al.*, 1995) was precipitated with 5% polyethylene glycol 3000 at 4°C and centrifuged at 20,000 xg for 10 min. The pellet was  
10 dissolved in 1 ml 20 mM Tris, pH 8.0 and 50 mM NaCl and pre-cleared by incubation with 200 µl protein A beads for 4 hrs at 4°C. Following incubation, 500 µl of pre-cleared brain supernatant was incubated with 30 µl of either anti-rsec8 antibody or rabbit immunoglobulin coupled to protein A beads overnight at 4°C. The beads were then washed three times with 200 µl 20 mM Tris, pH 8.0, 150 mM NaCl and 0.05% Tween 20. Proteins bound to the beads were solubilized in protein  
15 sample buffer, and subjected to Western blot analysis.

2. Immunoprecipitation of Synaptic Vesicles. Four frozen rat brains were powderized in liquid nitrogen using a Waring blender. The powderized brain was then homogenized in 20 ml of 0.3 M sucrose and 10 mM Hepes, pH 7.5 with a Teflon-glass homogenizer. The  
20 homogenate was centrifuged at 100,000 xg for 1 hr and the supernatant was incubated for 2 hrs at 4°C with 50 µl of anti-SV2 monoclonal antibody or purified mouse immunoglobulin coupled Dynabeads M-500 (Dynal, Oslo; 16 µg antibody/50 µl beads) in 0.3 M sucrose, 10 mM Hepes, pH 7.5 and 0.15% bovine serum albumin. Following incubation, the beads were washed four times with 0.3 M sucrose, 10 mM Hepes, pH 7.5 and solubilized in protein sample buffer. Protein  
25 samples were subjected to SDS-PAGE and analyzed by Western blotting.

3. Co-immunoprecipitation of syntaxin and the SA-17S complex. Anti-rsec8 monoclonal antibodies 2E9 and 17A10 or mouse immunoglobulin were coupled to protein G beads at a final concentration of 2 mg/ml using the cross-linker dimethylpimelimidate as previously described  
30 (Pevsner, *et al.*, 1994a). Rat brain membranes prepared as described above were resuspended in 20 mM Tris, pH 8.0, 150 mM NaCl and solubilized with either 1% CHAPS or Triton X-100 at a final protein concentration of 2 mg/ml. The solubilized sample was incubated at 4°C for 30 min before centrifuging at 20,000 xg for 20 min. Following centrifugation, the solubilized brain membranes

were pre-cleared by incubation with protein G beads (1 ml solubilized membranes per 200  $\mu$ l protein G beads) for 30 min at 4°C. The pre-cleared supernatant was then incubated with 20  $\mu$ l protein G beads with either immobilized anti-rsec8 monoclonal antibodies 2E9/17A10 or immobilized control mouse Ig for 4 hrs at 4°C. The beads were then washed three times with 300  $\mu$ l 20 mM Tris, pH 8.0, 150 mM NaCl containing either 0.7% CHAPS or Triton X-100. Proteins bound to the beads were analyzed by Western blotting.

#### G. Monoclonal and Polyclonal Antibodies

Mouse polyclonal antibodies against rsec6 and rsec8 were generated against a His-tagged rsec6 fusion protein. The mouse serum recognizes a single band of 86 kDa in brain homogenate. This band is not present when pre-immune serum was used as the primary antibody.

Mouse polyclonal antibodies against rsec8 were generated against a His-tagged rsec8 fusion protein. The immunized but not pre-immune mouse serum recognizes one single band at 110 kDa in brain homogenate.

Monoclonal antibodies against rsec8, including antibodies 2E9 and 17A10, were obtained from the corresponding hybridoma cell lines generated by fusion of NS-1 mouse myeloma cells with spleen cells from BALB/c mice immunized with His-tagged rsec8 fusion protein (Lane, *et al.*, 1986).

Rabbit polyclonal antibodies against rsec8 were generated and affinity-purified as described (Ting, *et al.*, 1995).

Polyclonal antibodies against nsec1 were affinity purified as described (Pevsner, *et al.*, 1994b).

Rabbit polyclonal antibodies against Na/K ATPase (Mays, *et al.*, 1995) was obtained from Dr. W. James Nelson. Monoclonal antibody against SV2 (Buckley and Kelly, 1985) was obtained from Dr. Kathy M. Buckley, and monoclonal antibody against synaptophysin was purchased from Boehringer Mannheim (Indianapolis, IN).

#### H. Hippocampal Cell Culture and Immunocytochemistry

Primary hippocampal CA3/CA1 cultures were obtained and maintained as described by Malgaroli and Tsien, 1992. Immunocytochemistry was carried out as previously described (Ting, *et al.*, 1995) using affinity-purified anti-rsec8 mouse polyclonal antibodies at 0.01  $\mu$ g/ml and affinity-purified rabbit anti-synaptotagmin rabbit polyclonal antibodies at 1:1000 dilution as primary antibodies. The labeling was visualized using fluorescein-labeled donkey anti-mouse and

Cy3-labeled donkey anti-rabbit antibodies. In control experiments, identical staining procedures were used except that the cells were labeled using pre-immune sera or purified immunoglobulin of equal protein concentration as the primary antibody instead.

5

#### EXAMPLE 1

##### rsec6 and rsec8 are Components of the Secretion-Associated 17S (SA-17S) Complex in Brain

To determine whether rsec6 and rsec8 function together in a protein complex, soluble brain proteins were fractionated on a linear glycerol gradient and the gradient fractions were analyzed by both sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and Western blotting.

10 A Coomassie blue-stained protein profile of the fractions is shown in Figure 1A. The positions of rsec8, rsec6 and nsec1 on the gel are indicated on the right of the figure as 8, 6, and 1 respectively. Most of the proteins fractionated between the 4.3S and 11.2S regions of the gradient.

The migration of rsec6 and rsec8 in the gradient was monitored by Western blot analysis using affinity-purified anti-rsec6 and anti-rsec8 antibodies. The results are shown in Figure 1B.

15 Fraction 1 corresponds to the top of the gradient. Size markers (indicated by the arrows) from left to right are bovine serum albumin (4.3S),  $\beta$ -amylase (11.2S) and thyroglobulin (19.2S). Both rsec6 and rsec8 migrated at the 17S position. In contrast, nsec1, another synaptosomal protein, migrated to the 4.3S position as previously observed (Pevsner, *et al.*, 1994b). This result indicates that all of the soluble rsec6 and rsec8 protein is in a large complex and that little or no protein is found in  
20 monomeric state.

Because the Coomassie blue-stained glycerol gradient protein profile (Figure 1A) showed the presence of many proteins in the 17S region, it was necessary to determine whether the co-fractionating rsec6 and rsec8 were in the same complex.

To address this question, rsec6 and rsec8 were immunoprecipitated from polyethylene  
25 glycol-fractionated rat brain supernatant with affinity-purified antibodies against rsec8 or with control rabbit immunoglobulin. The immunoprecipitated sample was analyzed for the presence of rsec6, rsec8 and nsec1 by Western blotting.

The results are shown in Figure 1C. Brain is the starting material; Ab represents proteins bound to the anti-rsec8 antibody in the absence of brain homogenate; Ab + brain represents proteins bound to the anti-rsec8 antibody following incubation with brain homogenate; and Ig + brain  
30 represents proteins bound to the control rabbit immunoglobulin following incubation with brain homogenate. The results show that rsec6, but not nsec1, co-immunoprecipitated with rsec8 from soluble brain proteins using affinity-purified rabbit polyclonal anti-rsec8 antibodies. No rsec6, rsec8

or rsec1 were immunoprecipitated by control rabbit immunoglobulin, confirming the specific binding of the SA-17S complex to anti-rsec8 antibodies. These data demonstrate that both rsec6 and rsec8 are components of the same high molecular weight complex.

5

## EXAMPLE 2

### Purification of the Mammalian SA-17S Complex

The SA-17S complex was purified from frozen rat brain by four sequential column chromatographic steps to define other members of the complex. Figure 2A shows a Coomassie blue-stained SDS polyacrylamide gel analysis of pooled column fractions. Abbreviations are as follows: Lys, starting soluble brain proteins; HT, hydroxyapatite eluant; TMAE, TMAE anion exchanger eluant; GF, HW55S gel filtration eluant; TMAE2, second TMAE anion exchanger eluant. Positions of rsec6 and rsec8 on the gel are indicated on the right as 6 and 8, respectively. Fifteen  $\mu$ g of proteins were loaded in lane Lys, HT and TMAE; eight  $\mu$ g of proteins were loaded in lane GF; and three  $\mu$ g of proteins were loaded in lane TMAE2.

The enrichment of the SA-17S complex following each chromatographic step was monitored by Western blot analysis using mouse polyclonal antibodies against rsec6 and rsec8, as shown in Figure 2B. Initially, cytosol prepared from frozen rat brains was applied to and step-eluted from hydroxyapatite. No significant amount of rsec6 and rsec8 was detected in the flow-through from this column. The pooled hydroxyapatite eluant exhibiting an enrichment of rsec6 and rsec8 immunoreactivity was concentrated and further purified using a strongly basic anion exchange resin, TMAE-Fractogel. All rsec6 and rsec8 immunoreactivity was retained by the resin and was eluted by a linear salt gradient. Next, taking advantage of the large size of the complex, the TMAE eluant was further fractionated by HW55-S gel filtration chromatography. The rsec6 and rsec8 immunoreactivity was eluted in a single peak with an apparent native molecular weight of 650 - 700 kDa. At this stage of purification, eight polypeptides of molecular weights between 70 kDa and 110 kDa were visible on a SDS polyacrylamide gel stained with Coomassie blue (Figure 2A).

The last stage of purification involved a second anion exchange chromatography utilizing TMAE-Fractogel. Analysis of salt-eluted fractions from this column by SDS-PAGE revealed that elution of eight distinct proteins with molecular weights ranging from 71 kDa to 110 kDa (Figure 2C) coincided exactly with the elution of rsec6 and rsec8 immunoreactivity. SDS-PAGE analysis of pooled peak rsec6 and rsec8 immunoreactivity fractions indicated that these same eight proteins constituted at least 95% of the total proteins in this fraction (Figure 2A, lane TMAE2; Figure 2C).

Fractionation of the purified complex on a glycerol gradient demonstrated that all eight components co-migrate at the 17S position. In addition, all eight proteins co-immunoprecipitated with rsec8 monoclonal antibodies. Approximately 100-200  $\mu$ g of SA-17S complex was recovered from 50 rat brains after purification. Quantitation of the enrichment of the SA-17S complex by Western blot using  $^{125}$ I-labeled secondary antibodies suggests that the complex constitutes approximately 0.1% of the total soluble brain protein. All eight proteins were approximately equally stained by Coomassie blue (although p79 reproducibly appears to stain somewhat less intensely) suggesting a stoichiometry of one copy of each protein within the complex (Figure 2C). This 1:1 stoichiometry is further supported by the apparent molecular weight of 600-700 kDa of the native complex. Peptide sequences from protein bands of 110 kDa and 86 kDa indicate that they are rsec8 and rsec6, respectively.

### EXAMPLE 3

#### Peptide Sequencing of the SA-17S Complex Components

To further characterize individual components of the mammalian SA-17S complex, the purified complex was fractionated on an SDS polyacrylamide gel and each polypeptide band was cut out and individually subjected to in-gel proteolysis with trypsin, Lys C and/or Glu C. The digested peptides were eluted from the gel, fractionated by high performance liquid chromatography (HPLC), and in many cases, analyzed by mass spectrometry for peptide purity and mass.

Figures 3A, 3B, 3C and 3D show the mass spectrometric analyzes of HPLC peptide peaks from four proteins, p102, p96, p84 and p71, respectively. A single peptide peak in the mass spectrometry profile suggests the presence of one peptide species in the analyzed HPLC peptide peak, as shown in Figs. 3A, 3B, 3C and 3D. The multiple peaks shown in Fig. 3D are due to the binding of different amounts of salt to the peptide (amino acid sequence analysis of this peptide peak confirmed the presence of only one peptide species). The mass of each peptide predicted by mass spectrometry is shown above the peptide peak. The amino acid sequence of each peptide determined by peptide sequencing and its mass (protonated form of peptide) calculated from the determined peptide sequence is shown beside corresponding peptide peak, along with its SEQ ID NO. In Panels A to C the calculated and observed masses matched with less than 0.1% difference.

Figures 4A-H, as well as Tables 1-8 below, show all the peptide sequences obtained from the eight predicted members of the SA-17S complex. Protein components of purified SA-17S complex were digested with a combination Trypsin (cuts after a lysine or an arginine residue), Lys C (cuts after a lysine residue) and/or Glu C (cuts after a glutamic acid) as described in the

Experimental Procedures. p110 or rsec8 peptides were obtained from a combination of Lys C and Trypsin digestion; p106 peptides 1-6 were obtained from a combination of Lys C and Trypsin digestion and p106 peptides 7-8 were obtained from Lys C digestion; p102 peptides 1-4 were obtained from a combination of Lys C and trypsin digestion and p102 peptides 5-7 were obtained from Lys C digestion; p96 peptides 1-3 were obtained from Lys C digestion and p96 peptides 4-6 were obtained from a combination of Lys C and Glu C digestion; p86 or rsec6 peptide 1 was obtained from Lys C digestion and p86 or rsec6 peptides 2-3 were obtained from a combination of Lys C and Glu C digestion; p84 peptide 1 was obtained from Lys C digestion and p84 peptides 2-9 were obtained from a combination of Lys C and Glu C digestion; p79 peptides 1-6 were obtained from Lys C digestion and p79 peptides 7-10 were obtained from a combination of Lys C and Glu C digestion; p71 peptides were obtained from Lys C digestion.

**Table 1 - rsec6**

Name	SEQ ID NO:	Sequence
rsec6 .1	5	DFRQSINTIEXL
rsec6 .2	6	QGPSQASPNYXP
rsec6 .3	7	AAIQSQLDGVRTGLSQ

**Table 2 - rsec8**

Name	SEQ ID NO:	Sequence
rsec8.1	8	XAPEGPLIDVXNI
rsec8.2	9	EFAAFFAK
rsec8.3	10	XLGVQRPLLQSTSIXE

**Table 3 - SA-17S p71**

Name	SEQ ID NO:	Sequence
p71.1	11	XNQVAFQHFQELDEHI
p71.2	12	VCHLXDQLEXVN
p71.3	13	LHLIAQELPFDRFSEVK

**Table 4 - SA-17S p79**

Name	SEQ ID NO:	Sequence
p79.1	14	XTDYIAE
p79.2	15	ETYGAFLSRSXG
p79.3	16	XXPPQGVYPNPASP
p79.4	17	XXQEEETLMFIRGN
p79.5	18	ALFIRDDXQF
p79.6	19	NLPVFQSCL
p79.7	20	AVEYFQDKFPD
p79.8	21	YRVEQVGDMIDRLFDTS
p79.9	22	VYEDPALSAIFLHNNYNY
p79.10	23	XXYGAFHLHRYSSVPFVYGXH

**Table 5 - SA-17S p84**

Name	SEQ ID NO:	Sequence
p84.1	24	QLSQQSDGDRDLQEWQRVQALAEETAQYK
p84.2	25	XLQLSFNFSEPNRQRP
p84.3	26	SIPLALLPAAAAGA
p84.4	27	DAVXQNSTQAAETEN
p84.5	28	DYRNDEA
p84.6	29	ENNPEEDDPS
p84.7	30	XLSQQSDXG
p84.8	31	AAALRAPPXVTS
p84.9	32	XKREPLE

**Table 6 - SA-17S p96**

Name	SEQ ID NO:	Sequence
p96.1	33	VVGQFPFQDTELEK
p96.2	34	XVYEIFDN
p96.3	35	DFLESIR
p96.4	36	XDQDLQLADYDHMT
p96.5	37	STNLLLTRLXN

**Table 7 - SA-17S p102**

Name	SEQ ID NO:	Sequence
p102.1	38	YLSGLQAPGXPASQSIGAQ
p102.2	39	GGLSTFFEAQDALSAIHQK
p102.3	40	ASNTADTXRQER
p102.4	41	XRENYIEGGI
p102.5	42	ENLGRLFENYI
p102.6	43	XDYDVVINDYE
p102.7	44	XIPXLSTRPANP

**Table 8 - SA-17S p106**

Name	SEQ ID NO:	Sequence
p106.1	45	XXYGEIAXK
p106.2	46	XATVSLPEK
p106.3	47	XDYGVIAND
p106.4	48	LIKYFFMVASVK
p106.5	49	ELPEFNLHFFK
p106.6	50	XLQDVDLASXR
p106.7	51	XNRXNEPAVNVL
p106.8	52	XQLXNIVEPEXIY

"X" represents residues whose identity has not been unambiguously determined.



As mentioned above, peptide sequences from p110 and p86 confirmed that these two proteins are rsec8 and rsec6 respectively. Furthermore, all of the peptide sequences from these two protein bands match the amino acid sequence predicted by rsec6 and rsec8 cDNAs demonstrating that the bands are free of contaminating proteins and that the SA-17S complex is substantially pure or  
5 substantially purified.

Comparison of remaining peptide sequences to "GENBANK" and a database of human expressed sequence tag (EST) cDNAs generated by the Institute for Genomic Research (TIGR) found matches to two proteins (p71 and p79). Overall, the amino acid sequences match the sequences predicted from the cDNAs with an accuracy greater than 97%. Taken together, these results suggest  
10 that members of the SA-17S complex are novel proteins which have not been previously characterized.

#### EXAMPLE 4

##### Expression Pattern of the SA-17S Complex

15 To determine the site of SA-17S complex function in brain, the brain regional distribution of these proteins was examined using SDS-PAGE/Western blot analyzes as described above (10  $\mu$ g of proteins loaded per lane). Figure 5A shows that rsec8 is expressed in all regions of rat brain examined and there is no significant brain regional variation. A similar expression pattern is also observed for rsec6, suggesting that both proteins are required by all brain cells.

20 To determine whether the SA-17S complex functions in multiple exocytotic pathways in addition to the regulated pathway underlying neurotransmission, the presence of rsec8 in cell lines derived from kidney, ovary, pituitary and adrenal medulla tissues was investigated with SDS-PAGE/Western blot analyzes as described above. The results are shown in Figure 5B. Affinity-purified anti-rsec8 rabbit polyclonal antibodies detected rsec8 as a 110 kDa protein in all the  
25 cell lines examined. rsec6 was also observed in all cell lines examined. The cells with the highest amounts of rsec8 were AtT-20 and PC-12 cells. Both of these cell lines have regulated secretory pathways and contain numerous secretory vesicles as well as synaptic proteins, indicating that the expression of the SA-17S complex may parallel the level of secretory activity in cells.

30

#### EXAMPLE 5

##### Subcellular Localization of the SA-17S Complex

Cell fractionation studies were carried out to determine the solubility properties of the SA-17S complex. Rat brain homogenate was subjected to a 1,000 xg spin to yield postnuclear supernatant

(PNS). The PNS was then subjected to either (i) a 100,000 xg spin to yield soluble (PNS sup) and membrane protein (PNS pellet) fractions or (ii) a 25,000 xg spin to yield a supernatant and a crude synaptosomal pellet (crude synaptosomes). The synaptosomal pellet was lysed in 2 mM Hepes, pH 7.4 and homogenized with a Polytron. The homogenate was then subjected to a 100,000 xg spin to yield both soluble (cytosol) and membrane (memb) fractions. The membrane pellets were resuspended in 20 mM Hepes, pH 7.4 to volumes equal to that of the corresponding supernatants. Five  $\mu$ l of each protein sample was loaded per lane.

Figure 6A shows results of a Western blot analysis for the presence of rsec6 and rsec8 in the subcellular fractionation of the SA-17S complex. Upon homogenization of freshly dissected rat brain, rsec6 and rsec8 were present in both soluble and insoluble fractions. Approximately 25% and 75% of rsec6 and rsec8 were found in the supernatant and the membrane pellet respectively. Similarly, in crude synaptosomes, the majority of rsec6 and rsec8 was found associated with synaptosomal membranes. Very little rsec6 and rsec8 were found in the synaptosomal cytosol. However, in frozen brain more rsec6 and rsec8 (approximately 50%) was found in the soluble fraction. The presence of rsec6 and rsec8 in the membrane fraction was not due to the entrapment of the complex in intact cells or organelles in the postnuclear supernatant because the distribution of these two proteins was similar when rat brain was completely homogenized in hypotonic buffer with a Polytron. Likewise, the presence of these two proteins in the supernatant was not due to contaminating membrane fragments because both SV2, a synaptic vesicle marker, and Na/K ATPase, a plasma membrane marker, were detected only in synaptosomal membrane pellet and not in synaptosomal cytosol.

To determine if the pool of the SA-17S complex that pelleted with crude synaptosomal membranes at 100,000 xg was due to its association with synaptosomal membranes or with large proteinaceous aggregates, crude synaptosomal membranes were subjected to sucrose gradient flotation. Crude synaptosomal membranes were placed at the bottom of a sucrose gradient and allowed to float up into the gradient. The flotation of the SA-17S complex into the gradient was monitored by Western blot analysis using mouse polyclonal antibodies. The migration of plasma membrane and synaptic vesicles was also monitored using rabbit polyclonal antibodies against Na/K ATPase and a monoclonal antibody against synaptophysin respectively. Fraction 1 corresponds to the top of the gradient.

The results are shown in Figure 6B. Both rsec6 and rsec8 floated up into the sucrose gradient, as observed with both Na/K ATPase, and synaptophysin (a synaptic vesicle marker). Very little rsec6 and rsec8 were detected at the bottom of the gradient as protein aggregates. These data

suggest that the insoluble pool of SA-17S complex is associated with synaptosomal (brain) membranes.

To investigate the nature of SA-17S complex association with membranes, brain membranes were extracted with 20 mM Tris buffer, 1.5 M NaCl, 4% Triton X-100, or 4 M urea. The  
5 extracted membranes were centrifuged and the resulting supernatants and membrane pellets were subjected to Western blot analyzes. Lysed brain membranes were extracted with either 20 mM Tris, pH 8.0, 1.5 M NaCl, 4% Triton X-100 or 4 M urea. Following extraction, the membranes were centrifuged to yield soluble (s) and membrane (p) fractions. The presence of rsec8 in these fractions was detected and quantitated by phosphorimaging using anti-rsec8 primary and <sup>125</sup>I-labeled secondary  
10 antibodies (see Materials and Methods, above).

The results are shown in Fig. 6C. Approximately 4%, 46%, 60% and 88% of the SA-17S complex were extracted by Tris buffer, 1.5 M NaCl, 4% Triton X-100 and 4 M urea, respectively. The incomplete extraction of the SA-17S complex in the presence of a high concentration of salt suggests that this complex may bind very tightly to the membranes. Interestingly, a fraction of the  
15 native SA-17S complex is found in the Triton X-100 insoluble pellet following membrane solubilization, contrary to the observation that purified soluble SA-17S complex is completely soluble in Triton X-100, suggesting that the membrane-bound complex may be associated with cytoskeletal elements. Taken together, these observations suggest that the insoluble pool of SA-17S complex may be associated with cytoskeletal elements which, in turn, bind to membranes.

Since rsec6 and rsec8 floated to the same region of the sucrose gradient as synaptic vesicles and plasma membranes, the possibility of an association between the SA-17S complex and synaptic vesicles was investigated. Synaptic vesicles were immunoprecipitated using an anti-SV2 (a synaptic vesicle-specific integral membrane protein) monoclonal antibody, or control mouse immunoglobulin, coupled to Dynabeads M-500. The immunoprecipitated vesicles were examined for the presence of  
20 synaptotagmin and rsec8 by Western blotting.

Figure 6D shows one such Western blot analysis. While the synaptic vesicle-specific protein synaptotagmin was present in anti-SV2 antibody immunoprecipitated vesicles, no rsec8 was detected even upon longer exposure. Neither protein was immunoprecipitated by control mouse immunoglobulin indicating a specific association of synaptic vesicles with anti-SV2 antibody. These  
30 results suggest that SA-17S complex is not associated strongly or stably with synaptic vesicles leaving the plasma membrane as the likely site of its membrane localization.

#### EXAMPLE 6

#### Localization of rsec8 in Cultured Hippocampal Neurons

To confirm the presence of the SA-17S complex in synapses, its localization in primary hippocampal cultures was studied by immunofluorescence microscopy. Fifteen to twenty day-old cultures of primary hippocampal neurons were fixed and immunostained with antibodies against rsec8 (Figs. 7A and 7B) and synaptotagmin (Fig. 7C). Fig. 7D shows a larger magnification of rsec8 staining overlaid with that of synaptotagmin at two synaptic terminals indicated by arrows. Synaptic terminals in these cultures were identified by the presence of the synaptic vesicle marker, synaptotagmin. Bar = 7 $\mu$ m.

Figure 7A shows that rsec8 labeling was detected throughout hippocampal neurons, including the cell bodies, dendrites, axons and nerve terminals. Brighter rsec8 staining at synapses suggests an accumulation of the SA-17S complex in nerve terminals (Figs. 7A and B). Although both synaptotagmin and rsec8 were detected in synaptic terminals, staining of rsec8 was more concentrated near the plasma membrane than that of synaptotagmin (Figs. 7B, 7C and 7D). Many terminals displayed rsec8 labeling in peripheral areas where it did not co-localize with the more centrally positioned synaptotagmin staining.

Glial and other non-neural cells in the hippocampal neuron cultures were also stained, although typically more weakly than the neurons. As controls, pre-immune immunoglobulins of same protein concentrations as the anti-rsec8 antibodies were used as the primary antibodies. At the same time exposures, no staining was observed in these neurons using identical labeling procedures. Taken together, these biochemical and histochemical data suggest that the SA-17S complex is localized to the plasma membrane.

#### EXAMPLE 7

##### Co-immunoprecipitation of rsec8 Complex with Syntaxin

Monoclonal antibodies were generated against rsec8 to study the properties of the SA-17S complex, including its membrane association in brain. The specificity of seven monoclonal antibodies generated against rsec8 was investigated by Western blot analysis as described above (ten  $\mu$ g of brain postnuclear supernatant were loaded per lane). As shown in Figure 8A, all these antibodies (with the exception of mAb15E12) recognized a single band of 110 kDa in the brain postnuclear supernatant. Monoclonal antibody 15E12 detected a major band of 110 kDa in addition to two minor bands of 65 kDa and 90 kDa. These two lower molecular weight bands are likely to be degradation products of rsec8.

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

5

(i) APPLICANT: Scheller, Richard H.

(ii) TITLE OF INVENTION: Methods and Compositions for Modulation  
of Vesicular Release

10

(iii) NUMBER OF SEQUENCES: 62

## (iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Dehlinger & Associates  
(B) STREET: 350 Cambridge Avenue, Suite 250  
(C) CITY: Palo Alto  
(D) STATE: CA  
(E) COUNTRY: USA  
(F) ZIP: 94306

20

## (v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Diskette  
(B) COMPUTER: IBM Compatible  
(C) OPERATING SYSTEM: DOS  
(D) SOFTWARE: FastSEQ for Windows Version 2.0

25

## (vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: To be assigned  
(B) FILING DATE: 19-DEC-1997  
(C) CLASSIFICATION:

30

## (vii) PRIOR APPLICATION DATA

(A) APPLICATION NUMBER: 60/033,905  
(B) FILING DATE: 20-DEC-1996

35

## (viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Dehlinger, Peter J.  
(B) REGISTRATION NUMBER: 28,006  
(C) REFERENCE/DOCKET NUMBER: 8600-0179.30

40

## (ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (650) 324-0880  
(B) TELEFAX: (650) 324-0960

45

## (2) INFORMATION FOR SEQ ID NO:1:

## (i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 2265 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA to mRNA

10

## (iii) HYPOTHETICAL: NO

## (iv) ANTI-SENSE: NO

## 15 (vi) ORIGINAL SOURCE:

(C) INDIVIDUAL ISOLATE: rsec6

## (ix) FEATURE:

- 20 (A) NAME/KEY: CDS  
 (B) LOCATION: 1..2265

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

25 ATG TGC AAA GAT TCT GCC TGC TTT TCG ACT ATG AAG GAG ACA GAC CTG  
 48  
 Met Cys Lys Asp Ser Ala Cys Phe Ser Thr Met Lys Glu Thr Asp Leu  
 1 5 10 15

30 GAG GCT GTT GCA ACA GCA GTC CAA AGG GTG GCT GGG ATG CTT CAG CGC  
 96  
 Glu Ala Val Ala Thr Ala Val Gln Arg Val Ala Gly Met Leu Gln Arg  
 20 25 30

35 CCA GAC CAG CTG GAC AAA GTG GAG CAG TAT CGC AGA AGG GAG GCT CGG 144  
 Pro Asp Gln Leu Asp Lys Val Glu Gln Tyr Arg Arg Arg Glu Ala Arg  
 35 40 45

AAG AAG GCC TCT GTG GAG GCC AGG CTA AAG GCC GCA ATC CAG TCT CAA 192  
 40 Lys Lys Ala Ser Val Glu Ala Arg Leu Lys Ala Ala Ile Gln Ser Gln  
 50 55 60

CTA GAT GGC GTC CGC ACA GGC CTA AGC CAA CTG CAC AAT GCA CTG AAT 240  
 Leu Asp Gly Val Arg Thr Gly Leu Ser Gln Leu His Asn Ala Leu Asn  
 45 65 70 75 80

38

	GAT GTC AAG GAT ATC CAG CAG TCA CTG GCT GAT GTG AGC AAG GAC TGG	288
	Asp Val Lys Asp Ile Gln Gln Ser Leu Ala Asp Val Ser Lys Asp Trp	
	85 90 95	
5	AGG CAG AGC ATC AAC ACC ATC GAG AGT CTC AAG GAC GTA AAA GAC GCA	336
	Arg Gln Ser Ile Asn Thr Ile Glu Ser Leu Lys Asp Val Lys Asp Ala	
	100 105 110	
	GTG GTG CAG CAC AGC CAG CTG GCT GCA GCT GTG GAG AAC CTC AAG AAC	384
10	Val Val Gln His Ser Gln Leu Ala Ala Val Glu Asn Leu Lys Asn	
	115 120 125	
	ATA TTC TCT GTG CCT GAG ATT GTG AGG GAG ACC CAA GAT CTC ATC GAG	432
	Ile Phe Ser Val Pro Glu Ile Val Arg Glu Thr Gln Asp Leu Ile Glu	
15	130 135 140	
	CAA GGG GCT CTC CTG CAG GCC CAC CGG AAG CTA ATG GAT CTG GAG TGC	480
	Gln Gly Ala Leu Leu Gln Ala His Arg Lys Leu Met Asp Leu Glu Cys	
	145 150 155 160	
20	TCC CGG GAT GGG CTA ATG TGT GAG CAG TAC CGC ATG GAC AGT GGG AAC	528
	Ser Arg Asp Gly Leu Met Cys Glu Gln Tyr Arg Met Asp Ser Gly Asn	
	165 170 175	
25	AAA CGG GAC ATG ACC CTC ATT CAT GGC TAC TTC GGT AGC ACA CAG GGG	576
	Lys Arg Asp Met Thr Leu Ile His Gly Tyr Phe Gly Ser Thr Gln Gly	
	180 185 190	
	CTC TCT GAC GAG CTG GCC AAG CAG CTG TGG ATG GTG CTG CAA AGA TCG	624
30	Leu Ser Asp Glu Leu Ala Lys Gln Leu Trp Met Val Leu Gln Arg Ser	
	195 200 205	
	CTG GTC ACT GTC CGC CGG GAT CCC ACC TTG CTG GTC TCC GTG GTC AGG	672
35	Leu Val Thr Val Arg Arg Asp Pro Thr Leu Leu Val Ser Val Val Arg	
	210 215 220	
	ATC ATT GAA AGG GAA GAA AAA ATT GAC AGG CGG ATA CTT GAT CGA AAA	720
	Ile Ile Glu Arg Glu Glu Lys Ile Asp Arg Arg Ile Leu Asp Arg Lys	
40	225 230 235 240	
	AAG CAA ACT GGG TTT GTT CCT CCT GGA AGG CCC AAA AAC TGG AAG GAA	768
	Lys Gln Thr Gly Phe Val Pro Pro Gly Arg Pro Lys Asn Trp Lys Glu	
	245 250 255	

45

39

	AAA ATG TTT GCC GTC TTG GAC AGA ACT GTG ACG ACT AGA ATC GAA GGC	816
	Lys Met Phe Ala Val Leu Asp Arg Thr Val Thr Thr Arg Ile Glu Gly	
	260 265 270	
5	ACG CAG GCG GAC ACC AGA GAA TCT GAC AAG ATG TGG CTC GTG CGC CAC	864
	Thr Gln Ala Asp Thr Arg Glu Ser Asp Lys Met Trp Leu Val Arg His	
	275 280 285	
	CTG GAG ATC ATT AGG AAG TAC GTC CTG GAT GAC CTC GTC ATC GCC AAG	912
10	Leu Glu Ile Ile Arg Lys Tyr Val Leu Asp Asp Leu Val Ile Ala Lys	
	290 295 300	
	AAC CTG CTG GTG CAG TGC TTC CCT CCT CAC TAT GAC ATC TTT AAG AAC	960
	Asn Leu Leu Val Gln Cys Phe Pro Pro His Tyr Asp Ile Phe Lys Asn	
15	305 310 315 320	
	CTC CTG AGC ATG TAC CAC CAG GCC TTG AGC ATT CGG ATG CAG GAC CTC	1008
	Leu Leu Ser Met Tyr His Gln Ala Leu Ser Ile Arg Met Gln Asp Leu	
	325 330 335	
20	GCC TCA GAG GAC CTT GAG GCC AAC GAG ATT GTG AGC CTC TTG ACC TGG	1056
	Ala Ser Glu Asp Leu Glu Ala Asn Glu Ile Val Ser Leu Leu Thr Trp	
	340 345 350	
25	GTC CTA AAT ACC TAC ACC AGT GCA GAG ATG ATG GGG AAT GTG GAG CTG	1104
	Val Leu Asn Thr Tyr Thr Ser Ala Glu Met Met Gly Asn Val Glu Leu	
	355 360 365	
30	GCC CCA GAG GTG GAC GTT AAT GCC CTG GAA CCG CTC CTC TCA CCA AAC	1152
	Ala Pro Glu Val Asp Val Asn Ala Leu Glu Pro Leu Leu Ser Pro Asn	
	370 375 380	
	GTG GTC TCT GAG CTG CTC GAC ACA TAC ATG TCA ACA CTC ACG TCC AAC	1200
35	Val Val Ser Glu Leu Leu Asp Thr Tyr Met Ser Thr Leu Thr Ser Asn	
	385 390 395 400	
	ATC ATT GCC TGG CTT CGG AAA GCA CTG GAG ACA GAC AAG AAA GAC TGG	1248
	Ile Ile Ala Trp Leu Arg Lys Ala Leu Glu Thr Asp Lys Lys Asp Trp	
40	405 410 415	
	AGC AAA GAG ACG GAG CCG GAA GCA GAC CAG GAC GGC TAT TAT CAG ACG	1296
	Ser Lys Glu Thr Glu Pro Glu Ala Asp Gln Asp Gly Tyr Tyr Gln Thr	
	420 425 430	

45



40

	ACA CTT CCT GCC ATT GTA TTC CAG ATG TTT GAA CAG AAT CTT CAA GTT	1344
	Thr Leu Pro Ala Ile Val Phe Gln Met Phe Glu Gln Asn Leu Gln Val	
	435 440 445	
5	GCT GCT CAA ATA AGT GAA GAT TTG AAA ACA AAG GTA TTA GTT TTG TGT	1392
	Ala Ala Gln Ile Ser Glu Asp Leu Lys Thr Lys Val Leu Val Leu Cys	
	450 455 460	
	CTT CAG CAA ATG AAT TCT TTC CTA AGT AGA TAC AAA GAG GAA GCC CAG	1440
10	Leu Gln Gln Met Asn Ser Phe Leu Ser Arg Tyr Lys Glu Glu Ala Gln	
	465 470 475 480	
	CTT TAC AAA GAG GAG CAC CTG AGA AAC CGG CAG CAC CCA CAC TGT TAC	1488
	Leu Tyr Lys Glu Glu His Leu Arg Asn Arg Gln His Pro His Cys Tyr	
15	485 490 495	
	GTG CAG TAC ATG GTC GCC ATC ATC AAT AAC TGC CAG ACC TTC AAA GAA	1536
	Val Gln Tyr Met Val Ala Ile Ile Asn Asn Cys Gln Thr Phe Lys Glu	
	500 505 510	
20	TCC ATC ATC AGT CTG AAA AGA AAA TAC CTA AAA CCC GAA ACA GAG GAA	1584
	Ser Ile Ile Ser Leu Lys Arg Lys Tyr Leu Lys Pro Glu Thr Glu Glu	
	515 520 525	
25	AGC CTG TGT CAG AGT CAG CCC AGC ATG GAT GGG ATT CTA GAT GCC ATT	1632
	Ser Leu Cys Gln Ser Gln Pro Ser Met Asp Gly Ile Leu Asp Ala Ile	
	530 535 540	
30	GCT AAG GAA GGC TGC AGC AGT CTG CTA GAG GAG GTC TTC CTG GAT CTA	1680
	Ala Lys Glu Gly Cys Ser Ser Leu Leu Glu Glu Val Phe Leu Asp Leu	
	545 550 555 560	
	GAG CAA CAT CTG AAT GAG CTA ATG ACA AAG AAG TGG ATG TTA GGA TCA	1728
35	Glu Gln His Leu Asn Glu Leu Met Thr Lys Lys Trp Met Leu Gly Ser	
	565 570 575	
	AAT GCT GTG GAC ATC ATC TGT GTC ACC GTA GAA GAC TAC TTC AAT GAC	1776
	Asn Ala Val Asp Ile Ile Cys Val Thr Val Glu Asp Tyr Phe Asn Asp	
40	580 585 590	
	TTC GCC AAA ATT AAA AAG CCA TAC AAA AAG AGG ATG ACA GCT GAG GCA	1824
	Phe Ala Lys Ile Lys Lys Pro Tyr Lys Lys Arg Met Thr Ala Glu Ala	
	595 600 605	

45

42

(A) LENGTH: 755 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

5 (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```

Met Cys Lys Asp Ser Ala Cys Phe Ser Thr Met Lys Glu Thr Asp Leu
10   1           5           10           15

Glu Ala Val Ala Thr Ala Val Gln Arg Val Ala Gly Met Leu Gln Arg
           20           25           30

15 Pro Asp Gln Leu Asp Lys Val Glu Gln Tyr Arg Arg Arg Glu Ala Arg
           35           40           45

Lys Lys Ala Ser Val Glu Ala Arg Leu Lys Ala Ala Ile Gln Ser Gln
           50           55           60
20
Leu Asp Gly Val Arg Thr Gly Leu Ser Gln Leu His Asn Ala Leu Asn
           65           70           75           80

Asp Val Lys Asp Ile Gln Gln Ser Leu Ala Asp Val Ser Lys Asp Trp
25           85           90           95

Arg Gln Ser Ile Asn Thr Ile Glu Ser Leu Lys Asp Val Lys Asp Ala
           100          105          110

30 Val Val Gln His Ser Gln Leu Ala Ala Ala Val Glu Asn Leu Lys Asn
           115          120          125

Ile Phe Ser Val Pro Glu Ile Val Arg Glu Thr Gln Asp Leu Ile Glu
           130          135          140
35
Gln Gly Ala Leu Leu Gln Ala His Arg Lys Leu Met Asp Leu Glu Cys
145          150          155          160

Ser Arg Asp Gly Leu Met Cys Glu Gln Tyr Arg Met Asp Ser Gly Asn
40           165          170          175

Lys Arg Asp Met Thr Leu Ile His Gly Tyr Phe Gly Ser Thr Gln Gly
           180          185          190

45 Leu Ser Asp Glu Leu Ala Lys Gln Leu Trp Met Val Leu Gln Arg Ser

```

43

	195	200	205
	Leu Val Thr Val Arg Arg Asp Pro Thr Leu Leu Val Ser Val Val Arg		
	210	215	220
5	Ile Ile Glu Arg Glu Glu Lys Ile Asp Arg Arg Ile Leu Asp Arg Lys		
	225	230	235 240
	Lys Gln Thr Gly Phe Val Pro Pro Gly Arg Pro Lys Asn Trp Lys Glu		
10	245	250	255
	Lys Met Phe Ala Val Leu Asp Arg Thr Val Thr Thr Arg Ile Glu Gly		
	260	265	270
15	Thr Gln Ala Asp Thr Arg Glu Ser Asp Lys Met Trp Leu Val Arg His		
	275	280	285
	Leu Glu Ile Ile Arg Lys Tyr Val Leu Asp Asp Leu Val Ile Ala Lys		
	290	295	300
20	Asn Leu Leu Val Gln Cys Phe Pro Pro His Tyr Asp Ile Phe Lys Asn		
	305	310	315 320
	Leu Leu Ser Met Tyr His Gln Ala Leu Ser Ile Arg Met Gln Asp Leu		
25	325	330	335
	Ala Ser Glu Asp Leu Glu Ala Asn Glu Ile Val Ser Leu Leu Thr Trp		
	340	345	350
30	Val Leu Asn Thr Tyr Thr Ser Ala Glu Met Met Gly Asn Val Glu Leu		
	355	360	365
	Ala Pro Glu Val Asp Val Asn Ala Leu Glu Pro Leu Leu Ser Pro Asn		
	370	375	380
35	Val Val Ser Glu Leu Leu Asp Thr Tyr Met Ser Thr Leu Thr Ser Asn		
	385	390	395 400
	Ile Ile Ala Trp Leu Arg Lys Ala Leu Glu Thr Asp Lys Lys Asp Trp		
40	405	410	415
	Ser Lys Glu Thr Glu Pro Glu Ala Asp Gln Asp Gly Tyr Tyr Gln Thr		
	420	425	430
45	Thr Leu Pro Ala Ile Val Phe Gln Met Phe Glu Gln Asn Leu Gln Val		

44

	435	440	445
	Ala Ala Gln Ile Ser Glu Asp Leu Lys Thr Lys Val Leu Val Leu Cys		
	450	455	460
5	Leu Gln Gln Met Asn Ser Phe Leu Ser Arg Tyr Lys Glu Glu Ala Gln		
	465	470	475 480
	Leu Tyr Lys Glu Glu His Leu Arg Asn Arg Gln His Pro His Cys Tyr		
10	485	490	495
	Val Gln Tyr Met Val Ala Ile Ile Asn Asn Cys Gln Thr Phe Lys Glu		
	500	505	510
15	Ser Ile Ile Ser Leu Lys Arg Lys Tyr Leu Lys Pro Glu Thr Glu Glu		
	515	520	525
	Ser Leu Cys Gln Ser Gln Pro Ser Met Asp Gly Ile Leu Asp Ala Ile		
	530	535	540
20	Ala Lys Glu Gly Cys Ser Ser Leu Leu Glu Glu Val Phe Leu Asp Leu		
	545	550	555 560
	Glu Gln His Leu Asn Glu Leu Met Thr Lys Lys Trp Met Leu Gly Ser		
25	565	570	575
	Asn Ala Val Asp Ile Ile Cys Val Thr Val Glu Asp Tyr Phe Asn Asp		
	580	585	590
30	Phe Ala Lys Ile Lys Lys Pro Tyr Lys Lys Arg Met Thr Ala Glu Ala		
	595	600	605
	His Arg Arg Val Val Val Glu Tyr Leu Arg Ala Val Met Gln Lys Arg		
	610	615	620
35	Ile Ser Phe Arg Ser Ala Glu Glu Arg Lys Glu Gly Ala Glu Lys Met		
	625	630	635 640
	Val Arg Glu Ala Glu Gln Leu Arg Phe Leu Phe Arg Lys Leu Ala Ser		
40	645	650	655
	Gly Phe Gly Glu Asp Ala Asp Gly His Cys Asp Thr Ile Val Ala Val		
	660	665	670
45	Ala Glu Val Ile Lys Leu Thr Asp Pro Ser Leu Leu Tyr Leu Glu Val		

45

675

680

685

Ser Thr Leu Val Ser Lys Tyr Pro Asp Ile Arg Asp Asp His Ile Gly

690

695

700

5

Ala Leu Leu Ala Leu Arg Gly Asp Ala Ser Arg Asp Met Lys Gln Thr

705

710

715

720

Ile Met Glu Thr Leu Glu Gln Gly Pro Met Gln Ala Ser Pro Asn Tyr

10

725

730

735

Val Pro Ile Phe Gln Glu Ile Val Val Pro Ser Leu Asn Val Ala Lys

740

745

750

15 Leu Leu Lys

755

## (2) INFORMATION FOR SEQ ID NO:3:

20

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2925 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

25

(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA to mRNA

## (iii) HYPOTHETICAL: NO

30

## (iv) ANTI-SENSE: NO

## (vi) ORIGINAL SOURCE:

(C) INDIVIDUAL ISOLATE: rsec8

35

## (ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 1..2925

40

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ATG GCG GCA GAA GCA GCT GGT GGG AAA TAC AGG AGC ACA GTC AGC AAA

48

45 Met Ala Ala Glu Ala Ala Gly Gly Lys Tyr Arg Ser Thr Val Ser Lys

46

	1	5	10	15	
	AGC AAA GAC CCC TCG GGG CTG CTC ATC TCG GTG ATC AGG ACT CTG TCT				
	96				
5	Ser Lys Asp Pro Ser Gly Leu Leu Ile Ser Val Ile Arg Thr Leu Ser				
	20	25	30		
	ACC AGT GAT GAT GTT GAA GAC CGA GAA AAT GAG AAA GGT CGC CTT GAA				144
	Thr Ser Asp Asp Val Glu Asp Arg Glu Asn Glu Lys Gly Arg Leu Glu				
10	35	40	45		
	GAA GCC TAT GAG AAG TGT GAC CGT GAC CTG GAT GAA TTG ATC GTT CAG				192
	Glu Ala Tyr Glu Lys Cys Asp Arg Asp Leu Asp Glu Leu Ile Val Gln				
	50	55	60		
15	CAC TAC ACA GAA TTG ACA ACA GCC ATT CGC ACA TAC CAG AGC ATC ACA				240
	His Tyr Thr Glu Leu Thr Thr Ala Ile Arg Thr Tyr Gln Ser Ile Thr				
	65	70	75	80	
20	GAA CGC ATC ACT AAC TCC AGG AAT AAA ATC AAG CAG GTA AAA GAG AAC				288
	Glu Arg Ile Thr Asn Ser Arg Asn Lys Ile Lys Gln Val Lys Glu Asn				
	85	90	95		
	CTG CTT TCC TGC AAG ATG CTG TTG CAT TGC AAA CGG GAT GAG CTT CGT				336
25	Leu Leu Ser Cys Lys Met Leu Leu His Cys Lys Arg Asp Glu Leu Arg				
	100	105	110		
	AAA CTA TGG ATT GAA GGA ATT GAG CAT AAG CAT GTC CTG AAC CTT CTG				384
	Lys Leu Trp Ile Glu Gly Ile Glu His Lys His Val Leu Asn Leu Leu				
30	115	120	125		
	GAT GAG ATT GAA AAC ATC AAG CAA GTG CCT CAA AAG CTG GAA CAG TGC				432
	Asp Glu Ile Glu Asn Ile Lys Gln Val Pro Gln Lys Leu Glu Gln Cys				
	130	135	140		
35	ATG GCC AGC AAG CAC TAC CTC AGC GCC ACG GAC ATG CTG GTG TCA GCA				480
	Met Ala Ser Lys His Tyr Leu Ser Ala Thr Asp Met Leu Val Ser Ala				
	145	150	155	160	
40	GTG GAG TCT CTG GAG GGT CCG CTG CTC CAG GTG GAA GGA CTC AGT GAC				528
	Val Glu Ser Leu Glu Gly Pro Leu Leu Gln Val Glu Gly Leu Ser Asp				
	165	170	175		
	CTC AGG CTG GAA CTT CAC AGC AAG AAG ATG AAC CTG CAC TTG GTT CTC				576
45	Leu Arg Leu Glu Leu His Ser Lys Lys Met Asn Leu His Leu Val Leu				

47

	180	185	190	
	ATA GAG GAA CTG CAC CGA CAT CTG TAC ATC AAA TCC ACT AGC CGG GTT			624
5	Ile Glu Glu Leu His Arg His Leu Tyr Ile Lys Ser Thr Ser Arg Val			
	195	200	205	
	GTA CAG CGT AAC AAG GAA AAA GGG AAG ATG AGT TCT CAT GGC AAA GAT			672
10	Val Gln Arg Asn Lys Glu Lys Gly Lys Met Ser Ser His Gly Lys Asp			
	210	215	220	
	GCT TCT CCT GGT CCT CTG ATT GAT GTT TCA AAC ATC TCT ACT CCA CGC			720
15	Ala Ser Pro Gly Pro Leu Ile Asp Val Ser Asn Ile Ser Thr Pro Arg			
	225	230	235	240
	AAA TTC CTT GAT GCC ACT CAA TAT TCT GCT GCT GGA AGC TCA AGT GTG			768
	Lys Phe Leu Asp Ala Thr Gln Tyr Ser Ala Ala Gly Ser Ser Ser Val			
	245	250	255	
20	AGG GAG ATG AAC CTG CAG GAC ATC AAA GAG GAC TTG GAC TGT GAT CCT			816
	Arg Glu Met Asn Leu Gln Asp Ile Lys Glu Asp Leu Asp Cys Asp Pro			
	260	265	270	
	GAG GAG AAC AGC ACT CTT TTC ATG GGA ATT CTT ATT CAG GGG CTG GCC			864
25	Glu Glu Asn Ser Thr Leu Phe Met Gly Ile Leu Ile Gln Gly Leu Ala			
	275	280	285	
	AGA CTG AAG AAG ATC CCT GAG ACC GTT AAA GCC ATT AAA GAG CGT TTG			912
30	Arg Leu Lys Lys Ile Pro Glu Thr Val Lys Ala Ile Lys Glu Arg Leu			
	290	295	300	
	GAG CAG GAG CTA AAG CAG ATT GTG AAG AGG TCA ACC ACC CAG GTG GCA			960
35	Glu Gln Glu Leu Lys Gln Ile Val Lys Arg Ser Thr Thr Gln Val Ala			
	305	310	315	320
	GAC AGT GCC TAT CAG AGG GGA GAG AGC CTC ACT GTG GAC AAC CAG CCA			1008
	Asp Ser Ala Tyr Gln Arg Gly Glu Ser Leu Thr Val Asp Asn Gln Pro			
	325	330	335	
40	AGG TTA CTT CTA GAA CTG CTG GAA CTG TTG TTT GAT AAG TTT AAT GCT			1056
	Arg Leu Leu Glu Leu Leu Glu Leu Phe Asp Lys Phe Asn Ala			
	340	345	350	
45	GTA GCC AGT GCA CAT TCT ATA GTC CTG GGG TAC CTT CAG GAC TCT GTG			1104

48

	Val Ala Ser Ala His Ser Ile Val Leu Gly Tyr Leu Gln Asp Ser Val	
	355 360 365	
5	GGG ACC CAG CCA ACA CAG CAG GAA GAG ATC AAA TTA TAC GAT ATG GCA Gly Thr Gln Pro Thr Gln Gln Glu Glu Ile Lys Leu Tyr Asp Met Ala	1152
	370 375 380	
10	GAT GTG TGG GTG AAG ATC CAG GAT GTG CTG CAG ATG CTG TTG ACT GAA Asp Val Trp Val Lys Ile Gln Asp Val Leu Gln Met Leu Leu Thr Glu	1200
	385 390 395 400	
15	TAC TTG GAT ATG AAG AAC ACA CGT ACT GCG TCA GAG CCG TCA GCT CAA Tyr Leu Asp Met Lys Asn Thr Arg Thr Ala Ser Glu Pro Ser Ala Gln	1248
	405 410 415	
	CTA AGC TAT GCC AGT ACT GGA CGA GAG TTC GCA GCC TTT TTT GCC AAG Leu Ser Tyr Ala Ser Thr Gly Arg Glu Phe Ala Ala Phe Phe Ala Lys	1296
	420 425 430	
20	AAG AAA CCA CAA AGG CCA AAG AAT TCT CTT TTC AAG TTT GAA TCA TCC Lys Lys Pro Gln Arg Pro Lys Asn Ser Leu Phe Lys Phe Glu Ser Ser	1344
	435 440 445	
25	TCC CAT GCT ATC AGT ATG AGC GCC TAT CTC CGA GAA CAG AGA AGG GAG Ser His Ala Ile Ser Met Ser Ala Tyr Leu Arg Glu Gln Arg Arg Glu	1392
	450 455 460	
30	CTG TAC AGT CGG AGT GGA GAA CTT CAA GGA GGT CCT GAT GAC AAC TTA Leu Tyr Ser Arg Ser Gly Glu Leu Gln Gly Gly Pro Asp Asp Asn Leu	1440
	465 470 475 480	
35	ATT GAA GGT GGA GGA ACA AAA TTT GTC TGC AAA CCT GGA GCC AGA AAT Ile Glu Gly Gly Gly Thr Lys Phe Val Cys Lys Pro Gly Ala Arg Asn	1488
	485 490 495	
	ATT ACC GTC ATA TTC CAT CCA CTG CTG AGA TTT ATT CAG GAG ATT GAG Ile Thr Val Ile Phe His Pro Leu Leu Arg Phe Ile Gln Glu Ile Glu	1536
	500 505 510	
40	CAT GCC CTG GGA CTT GGC CCT GCC AAA CAG TGT CTC CTT CGA GAG TTT His Ala Leu Gly Leu Gly Pro Ala Lys Gln Cys Leu Leu Arg Glu Phe	1584
	515 520 525	
45	CTC ACT ATA TAT ATC AAG AAC ATC TTC CTT AAT CAG GTC TTG ACT GAG	1632



	Leu Thr Ile Tyr Ile Lys Asn Ile Phe Leu Asn Gln Val Leu Thr Glu	
	530                      535                      540	
	ATC AAC AAG GAG ATC GAA GGT GTC ACC AAA ACC TCA GAC CCC TTG AAG	1680
5	Ile Asn Lys Glu Ile Glu Gly Val Thr Lys Thr Ser Asp Pro Leu Lys	
	545                      550                      555                      560	
	ATC CTA GCT AAT GCA GAC ACC ATG AAG GTT CTG GGC GTG CAG AGG CCT	1728
	Ile Leu Ala Asn Ala Asp Thr Met Lys Val Leu Gly Val Gln Arg Pro	
10	565                      570                      575	
	CTT CTA CAG AGC ACC ATC ATC GTG GAG AAG ACA GTA CAA GAC CTC ATG	1776
	Leu Leu Gln Ser Thr Ile Ile Val Glu Lys Thr Val Gln Asp Leu Met	
	580                      585                      590	
15	AAC CTG ATG CAT GAC TTG AGT GCA TAT TCA GAT CAG TTC CTC AAC ATG	1824
	Asn Leu Met His Asp Leu Ser Ala Tyr ser Asp Gln Phe Leu Asn Met	
	595                      600                      605	
20	GTG TGT GTG AAG CTA CAG GAA TAC AAG GAC ACA TGC TCC ACA GCC TAC	1872
	Val Cys Val Lys Leu Gln Glu Tyr Lys Asp Thr Cys Ser Thr Ala Tyr	
	610                      615                      620	
	AGA GGC ATT GTC CAA TCA GAA GAG AAA CTT GTC ATC AGT GCA TCT TGG	1920
25	Arg Gly Ile Val Gln Ser Glu Glu Lys Leu Val Ile Ser Ala Ser Trp	
	625                      630                      635                      640	
	GCA AAA GAT GAT GAT ATC AGC AGA CTC TTG AAA TCA CTG CCG AAC TGG	1968
	Ala Lys Asp Asp Asp Ile Ser Arg Leu Leu Lys Ser Leu Pro Asn Trp	
30	645                      650                      655	
	ACT AAC ATG GCT CAG CCC AAA CAG CTA AGG CCT AAG AGA GAG GAG GAG	2016
	Thr Asn Met Ala Gln Pro Lys Gln Leu Arg Pro Lys Arg Glu Glu Glu	
	660                      665                      670	
35		
	GAA GAC TTC ATA AGG GCA GCC TTT GGC AAG GAG TCT GAG GTT CTG ATT	2064
	Glu Asp Phe Ile Arg Ala Ala Phe Gly Lys Glu Ser Glu Val Leu Ile	
	675                      680                      685	
40	GGA AAC CTT GGT GAC AAA CTC ATT CCT CCA CAA GAT ATT CTC CGT GAC	2112
	Gly Asn Leu Gly Asp Lys Leu Ile Pro Pro Gln Asp Ile Leu Arg Asp	
	690                      695                      700	
45	GTC AGT GAC CTC AAA GCC TTG GCC AAC ATG CAT GAA AGC CTG GAA TGG	2160

	Val Ser Asp Leu Lys Ala Leu Ala Asn Met His Glu Ser Leu Glu Trp	
	705                      710                      715                      720	
	CTG GCA GGC CGA ACA AAG TCA GCT TTC TCC AGC CTT TCT GCA TCC CAG	2208
5	Leu Ala Gly Arg Thr Lys Ser Ala Phe Ser Ser Leu Ser Ala Ser Gln	
	725                      730                      735	
	ATG CTT TCC CCT GCT CAA GAG AGC CAC GTA AAC ATG GAC CTT CCT CCA	2256
	Met Leu Ser Pro Ala Gln Glu Ser His Val Asn Met Asp Leu Pro Pro	
10	740                      745                      750	
	GTG TCT GAA CAG ATC ATG CAG ACG CTG AGT GAA CTT GCC AAG TCC TTC	2304
	Val Ser Glu Gln Ile Met Gln Thr Leu Ser Glu Leu Ala Lys Ser Phe	
	755                      760                      765	
15	CAG GAC ATG GCT GAC CGC TGC TTG CTC GTC TTG CAT CTG GAA GTG AGA	2352
	Gln Asp Met Ala Asp Arg Cys Leu Leu Val Leu His Leu Glu Val Arg	
	770                      775                      780	
	GTT CAT TGT TTC CAC TAT CTC ATC CCT CTG GCA AAG GAG GGG AAC TAT	2400
20	Val His Cys Phe His Tyr Leu Ile Pro Leu Ala Lys Glu Gly Asn Tyr	
	785                      790                      795                      800	
	GCC ATT GTG GCC AAT GTG GAA AGC ATG GAT TAT GAC CCT TTA GTG GTC	2448
25	Ala Ile Val Ala Asn Val Glu Ser Met Asp Tyr Asp Pro Leu Val Val	
	805                      810                      815	
	AAG CTC AAC AAA GAC ATC AGT GCC ATG GAA GAG GCC ATG AGC GCC AGC	2496
	Lys Leu Asn Lys Asp Ile Ser Ala Met Glu Glu Ala Met Ser Ala Ser	
30	820                      825                      830	
	CTC CAG CAA CAC AAG TTC CAG TAC ATC TTC GAA GGT CTG GGA CAC CTC	2544
	Leu Gln Gln His Lys Phe Gln Tyr Ile Phe Glu Gly Leu Gly His Leu	
35	835                      840                      845	
	ATC TCC TGT ATC CTC ATT AAT GGG GCT CAG TAC TTC CGG CGC ATC AGT	2592
	Ile Ser Cys Ile Leu Ile Asn Gly Ala Gln Tyr Phe Arg Arg Ile Ser	
	850                      855                      860	
40	GAG TCT GGC ATC AAG AAA ATG TGT AGA AAC ATT TTT GTT CTT CAG CAG	2640
	Glu Ser Gly Ile Lys Lys Met Cys Arg Asn Ile Phe Val Leu Gln Gln	
	865                      870                      875                      880	
45	AAT TTG ACC AAC ATC ACC ATG TCA CGG GAG GCA GAC TTG GAC TTT GCA	2688

51

Asn Leu Thr Asn Ile Thr Met Ser Arg Glu Ala Asp Leu Asp Phe Ala  
 885 890 895

AGA CAG TAC TAC GAG ATG CTG TAC AAC ACT GCT GAT GAG CTC CTG AAC 2736  
 5 Arg Gln Tyr Tyr Glu Met Leu Tyr Asn Thr Ala Asp Glu Leu Leu Asn  
 900 905 910

CTG GTG GTG GAC CAG GGT GTG AAA TAC ACA GAG CTA GAG TAC ATC CAT 2784  
 10 Leu Val Val Asp Gln Gly Val Lys Tyr Thr Glu Leu Glu Tyr Ile His  
 915 920 925

GCC CTG ACC CTG CTG CAC CGA AGC CAG ACT GGC GTG GGG GAC CAG ACC 2832  
 15 Ala Leu Thr Leu Leu His Arg Ser Gln Thr Gly Val Gly Asp Gln Thr  
 930 935 940

ACC CAG AAC ACC AGG CTG CAG AGA CTC AAG GAG ATC ATC TGT GAG CAG 2880  
 Thr Gln Asn Thr Arg Leu Gln Arg Leu Lys Glu Ile Ile Cys Glu Gln  
 945 950 955 960

20 GCT GCC ATC AAG CAA GCC ACC AAG GAC AAG AAA ATA ACC ACT GTG 2925  
 Ala Ala Ile Lys Gln Ala Thr Lys Asp Lys Lys Ile Thr Thr Val  
 965 970 975

25 (2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 975 amino acids  
 (B) TYPE: amino acid  
 30 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

35 Met Ala Ala Glu Ala Ala Gly Gly Lys Tyr Arg Ser Thr Val Ser Lys  
 1 5 10 15

Ser Lys Asp Pro Ser Gly Leu Leu Ile Ser Val Ile Arg Thr Leu Ser  
 40 20 25 30

Thr Ser Asp Asp Val Glu Asp Arg Glu Asn Glu Lys Gly Arg Leu Glu  
 35 40 45

45 Glu Ala Tyr Glu Lys Cys Asp Arg Asp Leu Asp Glu Leu Ile Val Gln

52

	50		55		60
	His Tyr Thr Glu Leu Thr Thr Ala Ile Arg Thr Tyr Gln Ser Ile Thr				
	65		70		75 80
5	Glu Arg Ile Thr Asn Ser Arg Asn Lys Ile Lys Gln Val Lys Glu Asn				
		85		90	95
	Leu Leu Ser Cys Lys Met Leu Leu His Cys Lys Arg Asp Glu Leu Arg				
10		100		105	110
	Lys Leu Trp Ile Glu Gly Ile Glu His Lys His Val Leu Asn Leu Leu				
		115		120	125
15	Asp Glu Ile Glu Asn Ile Lys Gln Val Pro Gln Lys Leu Glu Gln Cys				
		130		135	140
	Met Ala Ser Lys His Tyr Leu Ser Ala Thr Asp Met Leu Val Ser Ala				
	145		150		155 160
20	Val Glu Ser Leu Glu Gly Pro Leu Leu Gln Val Glu Gly Leu Ser Asp				
		165		170	175
25	Leu Arg Leu Glu Leu His Ser Lys Lys Met Asn Leu His Leu Val Leu				
		180		185	190
	Ile Glu Glu Leu His Arg His Leu Tyr Ile Lys Ser Thr Ser Arg Val				
		195		200	205
30	Val Gln Arg Asn Lys Glu Lys Gly Lys Met Ser Ser His Gly Lys Asp				
		210		215	220
	Ala Ser Pro Gly Pro Leu Ile Asp Val Ser Asn Ile Ser Thr Pro Arg				
35	225		230		235 240
	Lys Phe Leu Asp Ala Thr Gln Tyr Ser Ala Ala Gly Ser Ser Ser Val				
		245		250	255
40	Arg Glu Met Asn Leu Gln Asp Ile Lys Glu Asp Leu Asp Cys Asp Pro				
		260		265	270
	Glu Glu Asn Ser Thr Leu Phe Met Gly Ile Leu Ile Gln Gly Leu Ala				
		275		280	285
45					

53

Arg Leu Lys Lys Ile Pro Glu Thr Val Lys Ala Ile Lys Glu Arg Leu  
 290 295 300

Glu Gln Glu Leu Lys Gln Ile Val Lys Arg Ser Thr Thr Gln Val Ala  
 5 305 310 315 320

Asp Ser Ala Tyr Gln Arg Gly Glu Ser Leu Thr Val Asp Asn Gln Pro  
 325 330 335

10 Arg Leu Leu Leu Glu Leu Leu Glu Leu Leu Phe Asp Lys Phe Asn Ala  
 340 345 350

Val Ala Ser Ala His Ser Ile Val Leu Gly Tyr Leu Gln Asp Ser Val  
 355 360 365

15 Gly Thr Gln Pro Thr Gln Gln Glu Glu Ile Lys Leu Tyr Asp Met Ala  
 370 375 380

Asp Val Trp Val Lys Ile Gln Asp Val Leu Gln Met Leu Leu Thr Glu  
 20 385 390 395 400

Tyr Leu Asp Met Lys Asn Thr Arg Thr Ala Ser Glu Pro Ser Ala Gln  
 405 410 415

25 Leu Ser Tyr Ala Ser Thr Gly Arg Glu Phe Ala Ala Phe Phe Ala Lys  
 420 425 430

Lys Lys Pro Gln Arg Pro Lys Asn Ser Leu Phe Lys Phe Glu Ser Ser  
 435 440 445

30 Ser His Ala Ile Ser Met Ser Ala Tyr Leu Arg Glu Gln Arg Arg Glu  
 450 455 460

Leu Tyr Ser Arg Ser Gly Glu Leu Gln Gly Gly Pro Asp Asp Asn Leu  
 35 465 470 475 480

Ile Glu Gly Gly Gly Thr Lys Phe Val Cys Lys Pro Gly Ala Arg Asn  
 485 490 495

40 Ile Thr Val Ile Phe His Pro Leu Leu Arg Phe Ile Gln Glu Ile Glu  
 500 505 510

His Ala Leu Gly Leu Gly Pro Ala Lys Gln Cys Leu Leu Arg Glu Phe  
 515 520 525

45

54

Leu Thr Ile Tyr Ile Lys Asn Ile Phe Leu Asn Gln Val Leu Thr Glu  
 530 535 540

Ile Asn Lys Glu Ile Glu Gly Val Thr Lys Thr Ser Asp Pro Leu Lys  
 5 545 550 555 560

Ile Leu Ala Asn Ala Asp Thr Met Lys Val Leu Gly Val Gln Arg Pro  
 565 570 575

10 Leu Leu Gln Ser Thr Ile Ile Val Glu Lys Thr Val Gln Asp Leu Met  
 580 585 590

Asn Leu Met His Asp Leu Ser Ala Tyr Ser Asp Gln Phe Leu Asn Met  
 595 600 605

15 Val Cys Val Lys Leu Gln Glu Tyr Lys Asp Thr Cys Ser Thr Ala Tyr  
 610 615 620

Arg Gly Ile Val Gln Ser Glu Glu Lys Leu Val Ile Ser Ala Ser Trp  
 20 625 630 635 640

Ala Lys Asp Asp Asp Ile Ser Arg Leu Leu Lys Scr Leu Pro Asn Trp  
 645 650 655

25 Thr Asn Met Ala Gln Pro Lys Gln Leu Arg Pro Lys Arg Glu Glu Glu  
 660 665 670

Glu Asp Phe Ile Arg Ala Ala Phe Gly Lys Glu Ser Glu Val Leu Ile  
 675 680 685

30 Gly Asn Leu Gly Asp Lys Leu Ile Pro Pro Gln Asp Ile Leu Arg Asp  
 690 695 700

Val Ser Asp Leu Lys Ala Leu Ala Asn Met His Glu Ser Leu Glu Trp  
 35 705 710 715 720

Leu Ala Gly Arg Thr Lys Ser Ala Phe Ser Ser Leu Ser Ala Ser Gln  
 725 730 735

40 Met Leu Ser Pro Ala Gln Glu Ser His Val Asn Met Asp Leu Pro Pro  
 740 745 750

Val Ser Glu Gln Ile Met Gln Thr Leu Ser Glu Leu Ala Lys Ser Phe  
 755 760 765

45

55

Gln Asp Met Ala Asp Arg Cys Leu Leu Val Leu His Leu Glu Val Arg  
 770 775 780  
 Val His Cys Phe His Tyr Leu Ile Pro Leu Ala Lys Glu Gly Asn Tyr  
 5 785 790 795 800  
 Ala Ile Val Ala Asn Val Glu Ser Met Asp Tyr Asp Pro Leu Val Val  
 805 810 815  
 10 Lys Leu Asn Lys Asp Ile Ser Ala Met Glu Glu Ala Met Ser Ala Ser  
 820 825 830  
 Leu Gln Gln His Lys Phe Gln Tyr Ile Phe Glu Gly Leu Gly His Leu  
 835 840 845  
 15 Ile Ser Cys Ile Leu Ile Asn Gly Ala Gln Tyr Phe Arg Arg Ile Ser  
 850 855 860  
 Glu Ser Gly Ile Lys Lys Met Cys Arg Asn Ile Phe Val Leu Gln Gln  
 20 865 870 875 880  
 Asn Leu Thr Asn Ile Thr Met Ser Arg Glu Ala Asp Leu Asp Phe Ala  
 885 890 895  
 25 Arg Gln Tyr Tyr Glu Met Leu Tyr Asn Thr Ala Asp Glu Leu Leu Asn  
 900 905 910  
 Leu Val Val Asp Gln Gly Val Lys Tyr Thr Glu Leu Glu Tyr Ile His  
 915 920 925  
 30 Ala Leu Thr Leu Leu His Arg Ser Gln Thr Gly Val Gly Asp Gln Thr  
 930 935 940  
 Thr Gln Asn Thr Arg Leu Gln Arg Leu Lys Glu Ile Ile Cys Glu Gln  
 35 945 950 955 960  
 Ala Ala Ile Lys Gln Ala Thr Lys Asp Lys Lys Ile Thr Thr Val  
 965 970 975

40

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 12 amino acids

45 (B) TYPE: amino acid

57

(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

5

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

10

(vi) ORIGINAL SOURCE:

(C) INDIVIDUAL ISOLATE: rsec6.3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

15

Ala Ala Ile Gln Ser Gln Leu Asp Gly Val Arg Thr Gly Leu Ser Gln  
1 5 10 15

20 (2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 13 amino acids

(B) TYPE: amino acid

25

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

30

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(C) INDIVIDUAL ISOLATE: rsec8.1

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Xaa Ala Pro Glu Gly Pro Leu Ile Asp Val Xaa Asn Ile

40

1

5

10

(2) INFORMATION FOR SEQ ID NO:9:

45

(i) SEQUENCE CHARACTERISTICS:



58

- (A) LENGTH: 8 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

- 5 (ii) MOLECULE TYPE: peptide
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- 10 (vi) ORIGINAL SOURCE:
  - (C) INDIVIDUAL ISOLATE: rsec8.2
- 15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Glu Phe Ala Ala Phe Phe Ala Lys  
1 5

- 20 (2) INFORMATION FOR SEQ ID NO:10:
  - (i) SEQUENCE CHARACTERISTICS:
    - 25 (A) LENGTH: 16 amino acids
    - (B) TYPE: amino acid
    - (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: peptide
  - 30 (iii) HYPOTHETICAL: NO
  - (iv) ANTI-SENSE: NO
  - (vi) ORIGINAL SOURCE:
    - 35 (C) INDIVIDUAL ISOLATE: rsec8.3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

40 Xaa Leu Gly Val Gln Arg Pro Leu Leu Gln Ser Thr Ser Ile Xaa Glu  
1 5 10 15

- (2) INFORMATION FOR SEQ ID NO:11:
- 45

60

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 17 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

5

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

10

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(C) INDIVIDUAL ISOLATE: p71.3

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Leu His Leu Ile Ala Gln Glu Leu Pro Phe Asp Arg Phe Ser Glu Val

1

5

10

15

20

Lys

## (2) INFORMATION FOR SEQ ID NO:14:

25

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 7 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

30

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

35

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(C) INDIVIDUAL ISOLATE: p79.1

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Xaa Thr Asp Tyr Ile Ala Glu

1

5

45

61

## (2) INFORMATION FOR SEQ ID NO:15:

## (i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 12 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## 10 (iii) HYPOTHETICAL: NO

## (iv) ANTI-SENSE: NO

## (vi) ORIGINAL SOURCE:

- 15 (C) INDIVIDUAL ISOLATE: p79.2

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

20 Glu Thr Tyr Gly Ala Phe Leu Ser Arg Ser Xaa Gly  
1 5 10

## (2) INFORMATION FOR SEQ ID NO:16:

25

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 14 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

30

## (ii) MOLECULE TYPE: peptide

## (iii) HYPOTHETICAL: NO

## 35 (iv) ANTI-SENSE: NO

## (vi) ORIGINAL SOURCE:

- (C) INDIVIDUAL ISOLATE: p79.3

40

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Xaa Xaa Pro Pro Gln Gly Val Pro Tyr Asn Pro Ala Ser Pro  
1 5 10

45

62

## (2) INFORMATION FOR SEQ ID NO:17:

## (i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 14 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## 10 (iii) HYPOTHETICAL: NO

## (iv) ANTI-SENSE: NO

## (vi) ORIGINAL SOURCE:

- 15 (C) INDIVIDUAL ISOLATE: p79.4

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

20 Xaa Xaa Gln Glu Glu Glu Thr Leu Met Phe Ile Arg Gly Asn  
1 5 10

## (2) INFORMATION FOR SEQ ID NO:18:

25

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

30

## (ii) MOLECULE TYPE: peptide

## (iii) HYPOTHETICAL: NO

## 35 (iv) ANTI-SENSE: NO

## (vi) ORIGINAL SOURCE:

- (C) INDIVIDUAL ISOLATE: p79.5

40

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Ala Leu Phe Ile Arg Asp Asp Xaa Gln Phe  
1 5 10

45

63

## (2) INFORMATION FOR SEQ ID NO:19:

## (i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 9 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## 10 (iii) HYPOTHETICAL: NO

## (iv) ANTI-SENSE: NO

## (vi) ORIGINAL SOURCE:

- 15 (C) INDIVIDUAL ISOLATE: p79.6

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

20 Asn Leu Pro Val Phe Gln Ser Cys Leu  
1 5

## (2) INFORMATION FOR SEQ ID NO:20:

25

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 11 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

30

## (ii) MOLECULE TYPE: peptide

## (iii) HYPOTHETICAL: NO

## 35 (iv) ANTI-SENSE: NO

## (vi) ORIGINAL SOURCE:

- (C) INDIVIDUAL ISOLATE: p79.7

40

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

Ala Val Glu Tyr Phe Gln Asp Lys Phe Pro Asp  
1 5 10

45

64

## (2) INFORMATION FOR SEQ ID NO:21:

## (i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 17 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## 10 (iii) HYPOTHETICAL: NO

## (iv) ANTI-SENSE: NO

## (vi) ORIGINAL SOURCE:

- 15 (C) INDIVIDUAL ISOLATE: p79.8

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

20 Tyr Arg Val Glu Gln Val Gly Asp Met Ile Asp Arg Leu Phe Asp Thr  
1 5 10 15  
Ser

## (2) INFORMATION FOR SEQ ID NO:22:

25

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

30

## (ii) MOLECULE TYPE: peptide

## (iii) HYPOTHETICAL: NO

## 35 (iv) ANTI-SENSE: NO

## (vi) ORIGINAL SOURCE:

- (C) INDIVIDUAL ISOLATE: p79.9

40

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Val Tyr Glu Asp Pro Ala Leu Ser Ala Ile Phe Leu His Asn Asn Tyr  
1 5 10 15

45

65

Asn Tyr

## (2) INFORMATION FOR SEQ ID NO:23:

5

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

10

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

15

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(C) INDIVIDUAL ISOLATE: p79.10

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

Xaa Xaa Tyr Gly Ala Phe Leu His Arg Tyr Ser Ser Val Pro Phe Val

1 5 10 15

25

Tyr Gly Xaa His

20

## 30 (2) INFORMATION FOR SEQ ID NO:24:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 30 amino acids

(B) TYPE: amino acid

35

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

40

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(C) INDIVIDUAL ISOLATE: p84.1

45

66

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Gln Leu Ser Gln Gln Ser Asp Gly Asp Arg Asp Leu Gln Glu Trp Arg  
1                      5                      10                      15  
5  
Gln Arg Val Gln Ala Leu Ala Glu Glu Thr Ala Gln Tyr Lys  
                    20                      25                      30

10 (2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 16 amino acids

(B) TYPE: amino acid

15 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

20

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(C) INDIVIDUAL ISOLATE: p84.2

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

Xaa Leu Gln Leu Ser Phe Asn Phe Ser Glu Pro Asn Arg Gln Arg Pro  
30 1                      5                      10                      15

(2) INFORMATION FOR SEQ ID NO:26:

35 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 14 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

40 (ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

45



67

(vi) ORIGINAL SOURCE:

(C) INDIVIDUAL ISOLATE: p84.3

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

Ser Ile Pro Leu Ala Leu Leu Pro Ala Ala Ala Gly Ala  
1 5 10

10 (2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 15 amino acids

(B) TYPE: amino acid

15 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

20

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(C) INDIVIDUAL ISOLATE: p84.4

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

Asp Ala Val Xaa Gln Asn Ser Thr Gln Ala Ala Glu Thr Glu Asn  
30 1 5 10 15

(2) INFORMATION FOR SEQ ID NO:28:

35 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 7 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

40 (ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

45

68

(vi) ORIGINAL SOURCE:

(C) INDIVIDUAL ISOLATE: p84.5

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

Asp Tyr Arg Asn Asp Glu Ala

1 5

10

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

15 (A) LENGTH: 10 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

20

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

25 (vi) ORIGINAL SOURCE:

(C) INDIVIDUAL ISOLATE: p84.6

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

30

Glu Asn Asn Pro Glu Glu Asp Asp Pro Ser

1 5 10

35 (2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 9 amino acids

(B) TYPE: amino acid

40 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

45

69

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(C) INDIVIDUAL ISOLATE: p84.7

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

Xaa Leu Ser Gln Gln Ser Asp Xaa Gly

10 1 5

(2) INFORMATION FOR SEQ ID NO:31:

15 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 12 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

20 (ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

25

(vi) ORIGINAL SOURCE:

(C) INDIVIDUAL ISOLATE: p84.8

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

Ala Ala Ala Leu Arg Ala Pro Pro Xaa Val Thr Ser

1 5 10

35

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 7 amino acids

40 (B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

45 (iii) HYPOTHETICAL: NO

70

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(C) INDIVIDUAL ISOLATE: p84.9

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

Xaa Lys Arg Glu Pro Leu Glu

10 1 5

(2) INFORMATION FOR SEQ ID NO:33:

15 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 14 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

20 (ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

25

(vi) ORIGINAL SOURCE:

(C) INDIVIDUAL ISOLATE: p96.1

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

Val Val Gly Gln Phe Pro Phe Gln Asp Thr Glu Leu Glu Lys

1 5 10

35 (2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 8 amino acids

(B) TYPE: amino acid

40 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

45

71

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(C) INDIVIDUAL ISOLATE: p96.2

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

Xaa Val Tyr Glu Ile Phe Asp Asn

10 1 5

(2) INFORMATION FOR SEQ ID NO:35:

15 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 7 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

20 (ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

25

(vi) ORIGINAL SOURCE:

(C) INDIVIDUAL ISOLATE: p96.3

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

Asp Phe Leu Glu Ser Ile Arg

1 5

35 (2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 14 amino acids

(B) TYPE: amino acid

40 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

45

72

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(C) INDIVIDUAL ISOLATE: p96.4

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

Xaa Asp Gln Asp Leu Gln Leu Ala Asp Tyr Asp His Met Thr  
10 1 5 10

(2) INFORMATION FOR SEQ ID NO:37:

15 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 12 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

20 (ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

25

(vi) ORIGINAL SOURCE:

(C) INDIVIDUAL ISOLATE: p96.5

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

Ser Thr Asn Leu Leu Leu Thr Arg Thr Leu Xaa Asn

1 5 10

35

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:

40 (A) LENGTH: 19 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

45

73

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

5 (vi) ORIGINAL SOURCE:

(C) INDIVIDUAL ISOLATE: p102.1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

10

Tyr Leu Ser Gly Leu Gln Ala Pro Gly Xaa Pro Ala Ser Gln Ser Ile  
1 5 10 15

Gly Ala Gln

15

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:

20

(A) LENGTH: 19 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

25

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

30 (vi) ORIGINAL SOURCE:

(C) INDIVIDUAL ISOLATE: p102.2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

35

Gly Gly Leu Ser Thr Phe Phe Glu Ala Gln Asp Ala Leu Ser Ala Ile  
1 5 10 15

His Gln Lys

40

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:

45

(A) LENGTH: 12 amino acids

74

(B) TYPE: amino acid  
(D) TOPOLOGY: linear

5 (ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

10 (vi) ORIGINAL SOURCE:

(C) INDIVIDUAL ISOLATE: p102.3

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

Ala Ser Asn Thr Ala Asp Thr Xaa Arg Gln Glu Arg  
1 5 10

20 (2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 amino acids

(B) TYPE: amino acid

25 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

30

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(C) INDIVIDUAL ISOLATE: p102.4

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

Xaa Arg Glu Asn Tyr Ile Glu Gly Gly Ile  
40 1 5 10

(2) INFORMATION FOR SEQ ID NO:42:

45 (i) SEQUENCE CHARACTERISTICS:



75

(A) LENGTH: 11 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

5 (ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

10

(vi) ORIGINAL SOURCE:

(C) INDIVIDUAL ISOLATE: p102.5

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

Glu Asn Leu Gly Arg Leu Phe Glu Asn Tyr Ile

1

5

10

20 (2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 11 amino acids

(B) TYPE: amino acid

25

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

30

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(C) INDIVIDUAL ISOLATE: p102.6

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

Xaa Asp Tyr Asp Val Val Ile Asn Asp Tyr Glu

40 1

5

10

(2) INFORMATION FOR SEQ ID NO:44:

45 (i) SEQUENCE CHARACTERISTICS:

76

- (A) LENGTH: 12 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

5       (ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

10

(vi) ORIGINAL SOURCE:

(C) INDIVIDUAL ISOLATE: p102.7

15       (xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

Xaa Ile Pro Xaa Leu Ser Thr Arg Pro Ala Asn Pro  
1                   5                   10

20

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 9 amino acids

25

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

30       (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

35       (C) INDIVIDUAL ISOLATE: p106.1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

40   Xaa Xaa Tyr Gly Glu Ile Ala Xaa Lys  
1                   5

(2) INFORMATION FOR SEQ ID NO:46:

45

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 9 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

5

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

10

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(C) INDIVIDUAL ISOLATE: p106.2

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

Xaa Ala Thr Val Ser Leu Pro Glu Lys

1

5

20 .

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:

25

(A) LENGTH: 9 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

30

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

35

(vi) ORIGINAL SOURCE:

(C) INDIVIDUAL ISOLATE: p106.3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

40

Xaa Asp Tyr Gly Val Ile Ala Asn Asp

1

5

45

(2) INFORMATION FOR SEQ ID NO:48:

78

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 12 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

5

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

10

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(C) INDIVIDUAL ISOLATE: p106.4

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

Leu Ile Lys Tyr Phe Phe Met Val Ala Ser Val Lys

1

5

10

20

## (2) INFORMATION FOR SEQ ID NO:49:

## (i) SEQUENCE CHARACTERISTICS:

25

(A) LENGTH: 11 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

30

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

35

(vi) ORIGINAL SOURCE:

(C) INDIVIDUAL ISOLATE: p106.5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

40

Glu Leu Pro Glu Phe Asn Leu His Phe Phe Lys

1

5

10

## (2) INFORMATION FOR SEQ ID NO:50:

45

79

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 11 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

5

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

10 (iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(C) INDIVIDUAL ISOLATE: p106.6

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

Xaa Leu Gln Asp Val Asp Leu Ala Ser Xaa Arg

1

5

10

20

## (2) INFORMATION FOR SEQ ID NO:51:

## (i) SEQUENCE CHARACTERISTICS:

25

(A) LENGTH: 12 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

30

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

35 (vi) ORIGINAL SOURCE:

(C) INDIVIDUAL ISOLATE: p106.7

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

40 Xaa Asn Arg Xaa Asn Glu Pro Ala Val Asn Val Leu

1

5

10

45

## (2) INFORMATION FOR SEQ ID NO:52:

80

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 13 amino acids  
    (B) TYPE: amino acid  
    (D) TOPOLOGY: linear
- 5           (ii) MOLECULE TYPE: peptide
- (iii) HYPOTHETICAL: NO
- 10           (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:  
                (C) INDIVIDUAL ISOLATE: p106.8
- 15           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:
- Xaa Gln Leu Xaa Asn Ile Val Glu Pro Glu Xaa Ile Tyr  
1                               5                               10
- 20
- (2) INFORMATION FOR SEQ ID NO:53:
- (i) SEQUENCE CHARACTERISTICS:  
25           (A) LENGTH: 2127 base pairs  
            (B) TYPE: nucleic acid  
            (C) STRANDEDNESS: single  
            (D) TOPOLOGY: linear
- 30           (ii) MOLECULE TYPE: cDNA to mRNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- 35           (vi) ORIGINAL SOURCE:  
                (C) INDIVIDUAL ISOLATE: p71 (rsec10)
- 40           (ix) FEATURE:  
            (A) NAME/KEY: Coding Sequence  
            (B) LOCATION: 1...2124
- 45           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

81

	ATG GCC ACG ACG GCC GAG CTC TTC GAG GAG CCT TTT GTG GCC GAT GAA	48
	Met Ala Thr Thr Ala Glu Leu Phe Glu Glu Pro Phe Val Ala Asp Glu	
	1                      5                      10                      15	
5	TAC ATT GAA CGT CTC GTG TGG AGA ACC CCG GGA GGT GGT TCT AGG GGT	96
	Tyr Ile Glu Arg Leu Val Trp Arg Thr Pro Gly Gly Gly Ser Arg Gly	
	20                      25                      30	
	GGA CCT GAA GCT TTT GAT CCT AAA AGG TTA TTA GAA GAA TTT GTG AAT	144
10	Gly Pro Glu Ala Phe Asp Pro Lys Arg Leu Leu Glu Glu Phe Val Asn	
	35                      40                      45	
	CAT ATT CAA GAA CTA CAG ATA ATG GAT GAA AGG ATT CAA AGG AAA GTA	192
	His Ile Gln Glu Leu Gln Ile Met Asp Glu Arg Ile Gln Arg Lys Val	
15	50                      55                      60	
	GAG AAG TTA GAA CAG CAG TGT CAG AAG GAG GCA AAG GAG TTT GCT AAG	240
	Glu Lys Leu Glu Gln Gln Cys Gln Lys Glu Ala Lys Glu Phe Ala Lys	
	65                      70                      75                      80	
20	AAG GTA CAA GAG CTG CAG AAA AGC AAC CAG GTT GCT TTC CAA CAT TTC	288
	Lys Val Gln Glu Leu Gln Lys Ser Asn Gln Val Ala Phe Gln His Phe	
	85                      90                      95	
25	CAA GAA CTA GAT GAA CAC ATT AGC TAT GTA GCT ACC AAA GTG TGT CAC	336
	Gln Glu Leu Asp Glu His Ile Ser Tyr Val Ala Thr Lys Val Cys His	
	100                      105                      110	
	CTT GGA GAC CAG TTG GAA GGG GTA AAC ACA CCT AGA CAG CGT GCA GTG	384
30	Leu Gly Asp Gln Leu Glu Gly Val Asn Thr Pro Arg Gln Arg Ala Val	
	115                      120                      125	
	GAG GCG CAG AAA TTG ATG AAA TAC TTT AAT GAG TTT CTA GAT GGA GAA	432
	Glu Ala Gln Lys Leu Met Lys Tyr Phe Asn Glu Phe Leu Asp Gly Glu	
35	130                      135                      140	
	TTG AAA TCT GAT GTT TTT ACA AAT CCT GAA AAG ATC AAG GAA GCA GCA	480
	Leu Lys Ser Asp Val Phe Thr Asn Pro Glu Lys Ile Lys Glu Ala Ala	
	145                      150                      155                      160	
40	GAC GTC ATT CAG AAA CTG CAT CTG ATC GCC CAG GAG TTA CCT TTC GAT	528
	Asp Val Ile Gln Lys Leu His Leu Ile Ala Gln Glu Leu Pro Phe Asp	
	165                      170                      175	
45	AGG TTT TCA GAA GTA AAA TCC AAA ATT GCA AGT AAA TAC CAC GAC TTA	576

82

	Arg Phe Ser Glu Val Lys Ser Lys Ile Ala Ser Lys Tyr His Asp Leu	
	180 185 190	
5	GAA TGC CAG CTG ATT CAG GAG TTT ACC AGT GCA CAA CGA AGA GGT GAG Glu Cys Gln Leu Ile Gln Glu Phe Thr Ser Ala Gln Arg Arg Gly Glu	624
	195 200 205	
10	GTC TCC CGG ATG AGG GAG GTG GCA GCT GTC CTG CTT CAT TTT AAG GGT Val Ser Arg Met Arg Glu Val Ala Ala Val Leu Leu His Phe Lys Gly	672
	210 215 220	
15	TAT TCC CAC TGT ATT GAT GTT TAT ATA AAG CAG TGC CAG GAG GGT GCT Tyr Ser His Cys Ile Asp Val Tyr Ile Lys Gln Cys Gln Glu Gly Ala	720
	225 230 235 240	
20	TAC TTG AGA AAT GAT ATA TTC GAA GAC GCA GCA ATT CTC TGT CAG CGA Tyr Leu Arg Asn Asp Ile Phe Glu Asp Ala Ala Ile Leu Cys Gln Arg	768
	245 250 255	
25	GTG AAC AAG CAA GTT GGA GAT ATC TTT AGT AAT CCA GAA GCA GTA CTG Val Asn Lys Gln Val Gly Asp Ile Phe Ser Asn Pro Glu Ala Val Leu	816
	260 265 270	
30	GCT AAG CTT ATT CAG AAT GTG TTT GAA GTC AAA CTA CAG AGT TTT GTG Ala Lys Leu Ile Gln Asn Val Phe Glu Val Lys Leu Gln Ser Phe Val	864
	275 280 285	
35	AAA GAT CAG TTA GAA GAA TGT AGG AAA TCG GAT GCT GAG CAG TAT CTT Lys Asp Gln Leu Glu Glu Cys Arg Lys Ser Asp Ala Glu Gln Tyr Leu	912
	290 295 300	
40	AAA AGT CTG TAT GAT CTG TAC ACA AGA ACC ACC AGT CTT TCC AGC AAA Lys Ser Leu Tyr Asp Leu Tyr Thr Arg Thr Ser Leu Ser Ser Lys	960
	305 310 315 320	
45	TTG ATG GAG TTT AAC TTA GGT ACT GAT AAA CAG ACT TTC TTG TCT AAG Leu Met Glu Phe Asn Leu Gly Thr Asp Lys Gln Thr Phe Leu Ser Lys	1008
	325 330 335	
50	CTT ATC AAA TCC ATT TTC GTT TCC TAT CTG GAG AAC TAT ATT GAA GTG Leu Ile Lys Ser Ile Phe Val Ser Tyr Leu Glu Asn Tyr Ile Glu Val	1056
	340 345 350	
55	GAG ATT GGA TAT TTG AAA AGC AGA AGT GCT ATG ATC CTA CAG CGC TAT	1104



83

	Glu Ile Gly Tyr Leu Lys Ser Arg Ser Ala Met Ile Leu Gln Arg Tyr	
	355 360 365	
5	TAT GAT TCA AAA AAC CAC CAA AAG AGA TCC ATT GGC ACA GGA GGT ATT Tyr Asp Ser Lys Asn His Gln Lys Arg Ser Ile Gly Thr Gly Gly Ile	1152
	370 375 380	
10	CAA GAT TTG AAA GAA AGG ATT AGA CAA CGT ACC AAC TTA CCA CTG GGG Gln Asp Leu Lys Glu Arg Ile Arg Gln Arg Thr Asn Leu Pro Leu Gly	1200
	385 390 395 400	
15	CCA AGT ATT GAC ACA CAT GGG GAG ACT TTC CTA TCT CAA GAA GTG GTA Pro Ser Ile Asp Thr His Gly Glu Thr Phe Leu Ser Gln Glu Val Val	1248
	405 410 415	
20	GTT AAT CTT TTA CAA GAA ACC AAA CAA GCC TTT GAA AGA TGT CAT AGG Val Asn Leu Leu Gln Glu Thr Lys Gln Ala Phe Glu Arg Cys His Arg	1296
	420 425 430	
25	CTT TCT GAT CCT TCT GAT TTA CCA AGG AAT GCC TTT AGA ATT TTT ACC Leu Ser Asp Pro Ser Asp Leu Pro Arg Asn Ala Phe Arg Ile Phe Thr	1344
	435 440 445	
30	ATT CTT GTG GAA TTT TTA TGT ATT GAG CAT ATT GAT TAT GCT TTA GAA Ile Leu Val Glu Phe Leu Cys Ile Glu His Ile Asp Tyr Ala Leu Glu	1392
	450 455 460	
35	ACG GGG CTT GCT GGA ATT CCA TCA TCA GAT TCT AGG AAC GCA AAT CTC Thr Gly Leu Ala Gly Ile Pro Ser Ser Asp Ser Arg Asn Ala Asn Leu	1440
	465 470 475 480	
40	TAT TTT TTG GAT GTT GTA CAA CAG GCC AAT ACT ATT TTT CAT CTT TTT Tyr Phe Leu Asp Val Val Gln Gln Ala Asn Thr Ile Phe His Leu Phe	1488
	485 490 495	
45	GAC AAA CAG TTT AAT GAT CAC CTA ATG CCA CTA ATA AGC TCT TCT CCT Asp Lys Gln Phe Asn Asp His Leu Met Pro Leu Ile Ser Ser Ser Pro	1536
	500 505 510	
50	AAG TTG TCT GAA TGC CTC CAG AAG AAG AAG GAG ATA ATT GAG CAG ATG Lys Leu Ser Glu Cys Leu Gln Lys Lys Lys Glu Ile Ile Glu Gln Met	1584
	515 520 525	
55	GAG ATG AAG TTG GAC ACT GGC ATC GAT AGG ACG TTA AAT TGT ATG ATT	1632

84

	Glu Met Lys Leu Asp Thr Gly Ile Asp Arg Thr Leu Asn Cys Met Ile	
	530 535 540	
5	GGA CAG ATG AAG CAC ATC TTG GCT GCA GAA CAA AAG AAA ACA GAT TTT Gly Gln Met Lys His Ile Leu Ala Ala Glu Gln Lys Lys Thr Asp Phe	1680
	545 550 555 560	
10	AAG CCA GAA GAT GAA AAC AAT GTT TTG ATT CAG TAT ACT AAT GCA TGT Lys Pro Glu Asp Glu Asn Asn Val Leu Ile Gln Tyr Thr Asn Ala Cys	1728
	565 570 575	
15	GTA AAG GTT TGT GCT TAT GTA AGA AAA CAA GTT GAG AAG ATT AAG AAT Val Lys Val Cys Ala Tyr Val Arg Lys Gln Val Glu Lys Ile Lys Asn	1776
	580 585 590	
20	TCC ATG GAT GGG AAG AAT GTG GAC ACG GTT CTG ATG GAG CTC GGA GTA Ser Met Asp Gly Lys Asn Val Asp Thr Val Leu Met Glu Leu Gly Val	1824
	595 600 605	
25	CGT TTC CAC CGG CTT ACT TAT GAG CAT CTT CAA CAA TAT TCC TAC AGT Arg Phe His Arg Leu Thr Tyr Glu His Leu Gln Gln Tyr Ser Tyr Ser	1872
	610 615 620	
30	TGT ATG GGC GGC ATG CTT GCA ATT TGC GAC GTT GCT GAA TAT AGA AAA Cys Met Gly Gly Met Leu Ala Ile Cys Asp Val Ala Glu Tyr Arg Lys	1920
	625 630 635 640	
35	TGT GCC AAA GAC TTC AAG ATT CCA ATG GTA TTA CAT CTT TTT GAT ACT Cys Ala Lys Asp Phe Lys Ile Pro Met Val Leu His Leu Phe Asp Thr	1968
	645 650 655	
40	CTG CAT GCA CTT TGC AAT CTC CTA GTC GTT GCT CCA GAT AAT TTA AAG Leu His Ala Leu Cys Asn Leu Leu Val Val Ala Pro Asp Asn Leu Lys	2016
	660 665 670	
45	CAA GTC TGC TCA GGA GAA CAA CTG GCC AAT CTG GAC AAG AAC ATA CTC Gln Val Cys Ser Gly Glu Gln Leu Ala Asn Leu Asp Lys Asn Ile Leu	2064
	675 680 685	
50	CAC TCC TTC GTC CAG CTT CGA GCT GAC TAT AGA TCC GCC CGC CTG GCT His Ser Phe Val Gln Leu Arg Ala Asp Tyr Arg Ser Ala Arg Leu Ala	2112
	690 695 700	
55	CGG CAC TTC AGC TGA	2127

85

Arg His Phe Ser

705

## 5 (2) INFORMATION FOR SEQ ID NO:54:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 708 amino acids

(B) TYPE: amino acid

10 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

15

Met Ala Thr Thr Ala Glu Leu Phe Glu Glu Pro Phe Val Ala Asp Glu  
 1 5 10 15

20

Tyr Ile Glu Arg Leu Val Trp Arg Thr Pro Gly Gly Gly Ser Arg Gly  
 20 25 30  
 Gly Pro Glu Ala Phe Asp Pro Lys Arg Leu Leu Glu Glu Phe Val Asn  
 35 40 45

25

His Ile Gln Glu Leu Gln Ile Met Asp Glu Arg Ile Gln Arg Lys Val  
 50 55 60  
 Glu Lys Leu Glu Gln Gln Cys Gln Lys Glu Ala Lys Glu Phe Ala Lys  
 65 70 75 80  
 Lys Val Gln Glu Leu Gln Lys Ser Asn Gln Val Ala Phe Gln His Phe  
 85 90 95

30

Gln Glu Leu Asp Glu His Ile Ser Tyr Val Ala Thr Lys Val Cys His  
 100 105 110  
 Leu Gly Asp Gln Leu Glu Gly Val Asn Thr Pro Arg Gln Arg Ala Val  
 115 120 125  
 Glu Ala Gln Lys Leu Met Lys Tyr Phe Asn Glu Phe Leu Asp Gly Glu  
 130 135 140

35

Leu Lys Ser Asp Val Phe Thr Asn Pro Glu Lys Ile Lys Glu Ala Ala  
 145 150 155 160  
 Asp Val Ile Gln Lys Leu His Leu Ile Ala Gln Glu Leu Pro Phe Asp  
 165 170 175

40

Arg Phe Ser Glu Val Lys Ser Lys Ile Ala Ser Lys Tyr His Asp Leu  
 180 185 190  
 Glu Cys Gln Leu Ile Gln Glu Phe Thr Ser Ala Gln Arg Arg Gly Glu  
 195 200 205  
 Val Ser Arg Met Arg Glu Val Ala Ala Val Leu Leu His Phe Lys Gly  
 210 215 220

45

Tyr Ser His Cys Ile Asp Val Tyr Ile Lys Gln Cys Gln Glu Gly Ala

86

	225		230		235		240									
	Tyr	Leu	Arg	Asn	Asp	Ile	Phe	Glu	Asp	Ala	Ala	Ile	Leu	Cys	Gln	Arg
				245					250					255		
	Val	Asn	Lys	Gln	Val	Gly	Asp	Ile	Phe	Ser	Asn	Pro	Glu	Ala	Val	Leu
5			260					265					270			
	Ala	Lys	Leu	Ile	Gln	Asn	Val	Phe	Glu	Val	Lys	Leu	Gln	Ser	Phe	Val
			275					280					285			
	Lys	Asp	Gln	Leu	Glu	Glu	Cys	Arg	Lys	Ser	Asp	Ala	Glu	Gln	Tyr	Leu
		290				295					300					
10	Lys	Ser	Leu	Tyr	Asp	Leu	Tyr	Thr	Arg	Thr	Thr	Ser	Leu	Ser	Ser	Lys
	305				310					315				320		
	Leu	Met	Glu	Phe	Asn	Leu	Gly	Thr	Asp	Lys	Gln	Thr	Phe	Leu	Ser	Lys
			325					330					335			
15	Leu	Ile	Lys	Ser	Ile	Phe	Val	Ser	Tyr	Leu	Glu	Asn	Tyr	Ile	Glu	Val
			340					345				350				
	Glu	Ile	Gly	Tyr	Leu	Lys	Ser	Arg	Ser	Ala	Met	Ile	Leu	Gln	Arg	Tyr
		355				360				365						
	Tyr	Asp	Ser	Lys	Asn	His	Gln	Lys	Arg	Ser	Ile	Gly	Thr	Gly	Gly	Ile
20		370				375				380						
	Gln	Asp	Leu	Lys	Glu	Arg	Ile	Arg	Gln	Arg	Thr	Asn	Leu	Pro	Leu	Gly
	385				390				395				400			
	Pro	Ser	Ile	Asp	Thr	His	Gly	Glu	Thr	Phe	Leu	Ser	Gln	Glu	Val	Val
			405					410				415				
25	Val	Asn	Leu	Leu	Gln	Glu	Thr	Lys	Gln	Ala	Phe	Glu	Arg	Cys	His	Arg
			420					425				430				
	Leu	Ser	Asp	Pro	Ser	Asp	Leu	Pro	Arg	Asn	Ala	Phe	Arg	Ile	Phe	Thr
		435				440				445						
	Ile	Leu	Val	Glu	Phe	Leu	Cys	Ile	Glu	His	Ile	Asp	Tyr	Ala	Leu	Glu
30		450				455				460						
	Thr	Gly	Leu	Ala	Gly	Ile	Pro	Ser	Ser	Asp	Ser	Arg	Asn	Ala	Asn	Leu
	465				470				475				480			
	Tyr	Phe	Leu	Asp	Val	Val	Gln	Gln	Ala	Asn	Thr	Ile	Phe	His	Leu	Phe
			485					490				495				
35	Asp	Lys	Gln	Phe	Asn	Asp	His	Leu	Met	Pro	Leu	Ile	Ser	Ser	Ser	Pro
		500				505				510						
	Lys	Leu	Ser	Glu	Cys	Leu	Gln	Lys	Lys	Lys	Glu	Ile	Ile	Glu	Gln	Met
		515				520				525						
	Glu	Met	Lys	Leu	Asp	Thr	Gly	Ile	Asp	Arg	Thr	Leu	Asn	Cys	Met	Ile
40		530				535				540						
	Gly	Gln	Met	Lys	His	Ile	Leu	Ala	Ala	Glu	Gln	Lys	Lys	Thr	Asp	Phe
	545				550				555				560			
	Lys	Pro	Glu	Asp	Glu	Asn	Asn	Val	Leu	Ile	Gln	Tyr	Thr	Asn	Ala	Cys
			565					570				575				
45	Val	Lys	Val	Cys	Ala	Tyr	Val	Arg	Lys	Gln	Val	Glu	Lys	Ile	Lys	Asn

87

580 585 590  
 Ser Met Asp Gly Lys Asn Val Asp Thr Val Leu Met Glu Leu Gly Val  
 595 600 605  
 Arg Phe His Arg Leu Thr Tyr Glu His Leu Gln Gln Tyr Ser Tyr Ser  
 5 610 615 620  
 Cys Met Gly Gly Met Leu Ala Ile Cys Asp Val Ala Glu Tyr Arg Lys  
 625 630 635 640  
 Cys Ala Lys Asp Phe Lys Ile Pro Met Val Leu His Leu Phe Asp Thr  
 645 650 655  
 10  
 Leu His Ala Leu Cys Asn Leu Leu Val Val Ala Pro Asp Asn Leu Lys  
 660 665 670  
 Gln Val Cys Ser Gly Glu Gln Leu Ala Asn Leu Asp Lys Asn Ile Leu  
 675 680 685  
 15 His Ser Phe Val Gln Leu Arg Ala Asp Tyr Arg Ser Ala Arg Leu Ala  
 690 695 700  
 Arg His Phe Ser  
 705

20

(2) INFORMATION FOR SEQ ID NO:55:

(i) SEQUENCE CHARACTERISTICS:

- 25 (A) LENGTH: 2328 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

30

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

35

(vi) ORIGINAL SOURCE:

(C) INDIVIDUAL ISOLATE: p79 (rexo70)

(ix) FEATURE:

- 40 (A) NAME/KEY: Coding Sequence  
 (B) LOCATION: 298...2256

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

88

	TCAGGTGTAG CTGCATGCTC TGAACACAGA CACTCATGTT CTTACCTGC TGAGGTAGCC	
60	AAGACATTTC CATCAGGCTC TTGGTGTGTC AAGCTTCTG CTGAGTCTGT GACATTTAGG	120
	TCCTGGACAT GAACTACCAT TGTGGGGACC TTTGGGAAGA GGCTCAAAGG GCTTATGTCT	180
5	TAAATCAATA TTTGTTTCCT TTTTGTCTGG TCACATTTGG GCTGTTGAG CTGCCATGTC	240
	TTGGAGCAGG GTTTGGAGCT AAGACATAGG GGAGTGGGAA CAGCGCTGTC CGCTGCG ATG	300
		Met
		1
10	ATT CCC CCG CAG GAG GCT TCC GCT CGG CGG CGG GAG ATC GAG GAC AAG	348
	Ile Pro Pro Gln Glu Ala Ser Ala Arg Arg Arg Glu Ile Glu Asp Lys	
	5 10 15	
	CTG AAG CAG GAG GAG GAG ACG CTG TCC TTT ATT CGA GAC AGC CTG GAG	396
15	Leu Lys Gln Glu Glu Glu Thr Leu Ser Phe Ile Arg Asp Ser Leu Glu	
	20 25 30	
	AAG AGC GAC CAG CTC ACT AAG AAC ATG GTG TCT ATC CTG TCG TCC TTT	444
	Lys Ser Asp Gln Leu Thr Lys Asn Met Val Ser Ile Leu Ser Ser Phe	
20	35 40 45	
	GAG AGC CGA CTC ATG AAG CTG GAG AAC TCC ATC ATC CCT GTG CAC AAG	492
	Glu Ser Arg Leu Met Lys Leu Glu Asn Ser Ile Ile Pro Val His Lys	
	50 55 60 65	
25	CAG ACA GAG AAC CTA CAG AGG CTG CAG GAG AAT GTC GAA AAG ACC TTA	540
	Gln Thr Glu Asn Leu Gln Arg Leu Gln Glu Asn Val Glu Lys Thr Leu	
	70 75 80	
30	TCC TGC CTG GAC CAT GTT ATC AGC TAT TAC CAC GTA GCC AGC GAC ACG	588
	Ser Cys Leu Asp His Val Ile Ser Tyr Tyr His Val Ala Ser Asp Thr	
	85 90 95	
	GAG AAG ATC ATC AGA GAG GGC CCC ACA GGT AGG CTG GAA GAA TAC CTG	636
35	Glu Lys Ile Ile Arg Glu Gly Pro Thr Gly Arg Leu Glu Glu Tyr Leu	
	100 105 110	
	GGA AGC ATG GCC AAA ATC CAG AAG GCT GTG GAG TAC TTT CAG GAC AAC	684
	Gly Ser Met Ala Lys Ile Gln Lys Ala Val Glu Tyr Phe Gln Asp Asn	
40	115 120 125	
	AGC CCA GAT AGC CCA GAG CTC AAC AAA GTG AAG CTG CTG TTT GAG CGG	732
	Ser Pro Asp Ser Pro Glu Leu Asn Lys Val Lys Leu Leu Phe Glu Arg	
	130 135 140 145	
45		

89

	GGG AAG GAG TCC TTG GAA TCG GAG TTC CGC AGT CTG ATG ACC AGG CAC	780
	Gly Lys Glu Ser Leu Glu Ser Glu Phe Arg Ser Leu Met Thr Arg His	
	150 155 160	
5	AGC AAG GTC ATC TCC CCT GTG CTC GTC CTG GAC CTC ATC AGT GCG GAT	828
	Ser Lys Val Ile Ser Pro Val Leu Val Leu Asp Leu Ile Ser Ala Asp	
	165 170 175	
10	GAC GAG CTG GAG GTC CAG GAG GAC GTG GTC CTG GAG CAC CTG CCT GAG	876
	Asp Glu Leu Glu Val Gln Glu Asp Val Val Leu Glu His Leu Pro Glu	
	180 185 190	
15	AGC GTG CTC CAG GAT GTG ATC CGC ATC TCC CGC TGG CTG GTG GAA TAT	924
	Ser Val Leu Gln Asp Val Ile Arg Ile Ser Arg Trp Leu Val Glu Tyr	
	195 200 205	
20	GGT CGG AAC CAA GAT TTC ATG AAT GTC TAC TAT CAG ATC CGC TCT AGC	972
	Gly Arg Asn Gln Asp Phe Met Asn Val Tyr Tyr Gln Ile Arg Ser Ser	
	210 215 220 225	
20	CAG CTG GAC CGC TCC ATC AAG GGT CTG AAG GAG CAT TTT CGC AAA AGC	1020
	Gln Leu Asp Arg Ser Ile Lys Gly Leu Lys Glu His Phe Arg Lys Ser	
	230 235 240	
25	AGT TCT TCC TCT GGG GTT CCC TAC TCC CCC GCC ATC CCC AAC AAG AGG	1068
	Ser Ser Ser Ser Gly Val Pro Tyr Ser Pro Ala Ile Pro Asn Lys Arg	
	245 250 255	
30	AAA GAC ACG CCT ACC AAG AAG CCC ATC AAG CGG CCA GGG AGA GAT GAC	1116
	Lys Asp Thr Pro Thr Lys Lys Pro Ile Lys Arg Pro Gly Arg Asp Asp	
	260 265 270	
35	ATG CTG GAC GTG GAG ACA GAT GCC TAC ATT CAC TGC GTT AGT GCC TTT	1164
	Met Leu Asp Val Glu Thr Asp Ala Tyr Ile His Cys Val Ser Ala Phe	
	275 280 285	
40	GTC AGG CTG GCA CAG AGT GAA TAC CAG CTG CTG ATG GGC ATC ATT CCC	1212
	Val Arg Leu Ala Gln Ser Glu Tyr Gln Leu Leu Met Gly Ile Ile Pro	
	290 295 300 305	
40	GAG CAT CAC CAG AAG AAA ACC TTC GAC TCC TTG ATA CAG GAT GCC CTA	1260
	Glu His His Gln Lys Lys Thr Phe Asp Ser Leu Ile Gln Asp Ala Leu	
	310 315 320	
45	GAT GGG CTG ATG CTT GAA GGG GAG AAC ATA GTG TCC GCG GCC AGG AAA	1308

90

Asp Gly Leu Met Leu Glu Gly Glu Asn Ile Val Ser Ala Ala Arg Lys  
 325 330 335

5 GCC ATC ATC CGC CAT GAC TTC TCC ACC GTG CTC ACC GTC TTC CCC ATC 1356  
 Ala Ile Ile Arg His Asp Phe Ser Thr Val Leu Thr Val Phe Pro Ile  
 340 345 350

10 CTG CGA CAC CTC AAG CAG ACC AAG CCT GAG TTT GAC CAG GTG CTC CAG 1404  
 Leu Arg His Leu Lys Gln Thr Lys Pro Glu Phe Asp Gln Val Leu Gln  
 355 360 365

15 GGC ACA GCA GCC AGC ACG AAG AAC AAG CTG CCA GGC CTC ATC ACC TCC 1452  
 Gly Thr Ala Ala Ser Thr Lys Asn Lys Leu Pro Gly Leu Ile Thr Ser  
 370 375 380 385

20 ATG GAG ACC ATT GGG GCC AAA GCT CTA GAA GAC TTT GCG GAT AAC ATC 1500  
 Met Glu Thr Ile Gly Ala Lys Ala Leu Glu Asp Phe Ala Asp Asn Ile  
 390 395 400

25 AAG AAT GAT CCA GAC AAG GAG TAC AAC ATG CCT AAA GAT GGC ACC GTT 1548  
 Lys Asn Asp Pro Asp Lys Glu Tyr Asn Met Pro Lys Asp Gly Thr Val  
 405 410 415

30 CAT GAG CTC ACA AGC AAT GCC ATC CTG TTC CTA CAG CAG CTT CTG GAC 1596  
 His Glu Leu Thr Ser Asn Ala Ile Leu Phe Leu Gln Gln Leu Leu Asp  
 420 425 430

35 TTC CAG GAG ACA GCA GGC GCC ATG CTG GCC TCC CAA GAA ACC AGC TCT 1644  
 Phe Gln Glu Thr Ala Gly Ala Met Leu Ala Ser Gln Glu Thr Ser Ser  
 435 440 445

40 TCG GCC ACC AGC TAC AAC TCC GAG TTC AGC AAG CGA CTG CTG AGT ACC 1692  
 Ser Ala Thr Ser Tyr Asn Ser Glu Phe Ser Lys Arg Leu Leu Ser Thr  
 450 455 460 465

45 TAC ATT TGC AAA GTC CTG GGT AAC CTG CAG CTG AAC TTG CTA AGC AAG 1740  
 Tyr Ile Cys Lys Val Leu Gly Asn Leu Gln Leu Asn Leu Leu Ser Lys  
 470 475 480

TCC AAG GTG TAT GAG GAT CCA GCT CTG AGC GCC ATC TTC CTA CAC AAC 1788  
 Ser Lys Val Tyr Glu Asp Pro Ala Leu Ser Ala Ile Phe Leu His Asn  
 485 490 495



AAC TAC AAC TAC ATC CTC AAG TCC CTG GAG AAG TCT GAG CTG ATC CAG 1836  
 Asn Tyr Asn Tyr Ile Leu Lys Ser Leu Glu Lys Ser Glu Leu Ile Gln  
 500 505 510

5 CTT GTG GCT GTG ACC CAG AAG ACT GCT GAG CGC TCC TAC CGG GAG CAC 1884  
 Leu Val Ala Val Thr Gln Lys Thr Ala Glu Arg Ser Tyr Arg Glu His  
 515 520 525

10 ATT GAG CAG CAG ATC CAG ACC TAC CAG CGC AGC TGG CTA AAG GTG ACT 1932  
 Ile Glu Gln Gln Ile Gln Thr Tyr Gln Arg Ser Trp Leu Lys Val Thr  
 530 535 540 545

15 GAC TAC ATC GCA GAG AAG AAT CTA CCT GTG TTC CAG CCT GGA GTC AAG 1980  
 Asp Tyr Ile Ala Glu Lys Asn Leu Pro Val Phe Gln Pro Gly Val Lys  
 550 555 560

20 CTC CGG GAC AAG GAG CGG CAA ATG ATT AAG GAA CGT TTC AAG GGA TTC 2028  
 Leu Arg Asp Lys Glu Arg Gln Met Ile Lys Glu Arg Phe Lys Gly Phe  
 565 570 575

AAT GAT GGG CTC GAA GAA CTG TGC AAG ATT CAG AAG GCC TGG GCC ATC 2076  
 Asn Asp Gly Leu Glu Glu Leu Cys Lys Ile Gln Lys Ala Trp Ala Ile  
 580 585 590

25 CCA GAC ACG GAG CAG AGG GAC AAG ATC CGA CAG GCT CAG AAA AGC ATC 2124  
 Pro Asp Thr Glu Gln Arg Asp Lys Ile Arg Gln Ala Gln Lys Ser Ile  
 595 600 605

30 GTC AAG GAG ACC TAT GGG GCC TTT CTG CAC AGG TAT AGC AGC GTG CCC 2172  
 Val Lys Glu Thr Tyr Gly Ala Phe Leu His Arg Tyr Ser Ser Val Pro  
 610 615 620 625

35 TTC ACC AAA AAC CCG GAG AAG TAC ATC AAG TAC CGT GTG GAG CAG GTG 2220  
 Phe Thr Lys Asn Pro Glu Lys Tyr Ile Lys Tyr Arg Val Glu Gln Val  
 630 635 640

GGC GAC ATG ATC GAC CGC CTC TTC GAC ACC TCT GCT TAAGTCGGCC AGCGCT 2272  
 Gly Asp Met Ile Asp Arg Leu Phe Asp Thr Ser Ala  
 645 650

45 TCCGCCAGT GCCACTGGCG TGCCTAGCAT GTCATCGGAC TCATAAACCA GTGGTA 2328

45 (2) INFORMATION FOR SEQ ID NO:56:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 653 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

5

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

```

10  Met Ile Pro Pro Gln Glu Ala Ser Ala Arg Arg Arg Glu Ile Glu Asp
    1             5             10             15
    Lys Leu Lys Gln Glu Glu Thr Leu Ser Phe Ile Arg Asp Ser Leu
        20             25             30
    Glu Lys Ser Asp Gln Leu Thr Lys Asn Met Val Ser Ile Leu Ser Ser
15      35             40             45
    Phe Glu Ser Arg Leu Met Lys Leu Glu Asn Ser Ile Ile Pro Val His
        50             55             60
    Lys Gln Thr Glu Asn Leu Gln Arg Leu Gln Glu Asn Val Glu Lys Thr
        65             70             75             80
20  Leu Ser Cys Leu Asp His Val Ile Ser Tyr Tyr His Val Ala Ser Asp
        85             90             95
    Thr Glu Lys Ile Ile Arg Glu Gly Pro Thr Gly Arg Leu Glu Glu Tyr
        100            105            110
    Leu Gly Ser Met Ala Lys Ile Gln Lys Ala Val Glu Tyr Phe Gln Asp
25      115            120            125
    Asn Ser Pro Asp Ser Pro Glu Leu Asn Lys Val Lys Leu Leu Phe Glu
        130            135            140
    Arg Gly Lys Glu Ser Leu Glu Ser Glu Phe Arg Ser Leu Met Thr Arg
        145            150            155            160
30  His Ser Lys Val Ile Ser Pro Val Leu Val Leu Asp Leu Ile Ser Ala
        165            170            175
    Asp Asp Glu Leu Glu Val Gln Glu Asp Val Val Leu Glu His Leu Pro
        180            185            190
    Glu Ser Val Leu Gln Asp Val Ile Arg Ile Ser Arg Trp Leu Val Glu
35      195            200            205
    Tyr Gly Arg Asn Gln Asp Phe Met Asn Val Tyr Tyr Gln Ile Arg Ser
        210            215            220
    Ser Gln Leu Asp Arg Ser Ile Lys Gly Leu Lys Glu His Phe Arg Lys
        225            230            235            240
40  Ser Ser Ser Ser Ser Gly Val Pro Tyr Ser Pro Ala Ile Pro Asn Lys
        245            250            255
    Arg Lys Asp Thr Pro Thr Lys Lys Pro Ile Lys Arg Pro Gly Arg Asp
        260            265            270
    Asp Met Leu Asp Val Glu Thr Asp Ala Tyr Ile His Cys Val Ser Ala
45      275            280            285

```

Phe Val Arg Leu Ala Gln Ser Glu Tyr Gln Leu Leu Met Gly Ile Ile  
 290 295 300  
 Pro Glu His His Gln Lys Lys Thr Phe Asp Ser Leu Ile Gln Asp Ala  
 305 310 315 320  
 5 Leu Asp Gly Leu Met Leu Glu Gly Glu Asn Ile Val Ser Ala Ala Arg  
 325 330 335  
 Lys Ala Ile Ile Arg His Asp Phe Ser Thr Val Leu Thr Val Phe Pro  
 340 345 350  
 Ile Leu Arg His Leu Lys Gln Thr Lys Pro Glu Phe Asp Gln Val Leu  
 10 355 360 365  
 Gln Gly Thr Ala Ala Ser Thr Lys Asn Lys Leu Pro Gly Leu Ile Thr  
 370 375 380  
 Ser Met Glu Thr Ile Gly Ala Lys Ala Leu Glu Asp Phe Ala Asp Asn  
 385 390 395 400  
 15 Ile Lys Asn Asp Pro Asp Lys Glu Tyr Asn Met Pro Lys Asp Gly Thr  
 405 410 415  
 Val His Glu Leu Thr Ser Asn Ala Ile Leu Phe Leu Gln Gln Leu Leu  
 420 425 430  
 Asp Phe Gln Glu Thr Ala Gly Ala Met Leu Ala Ser Gln Glu Thr Ser  
 20 435 440 445  
  
 Ser Ser Ala Thr Ser Tyr Asn Ser Glu Phe Ser Lys Arg Leu Leu Ser  
 450 455 460  
 Thr Tyr Ile Cys Lys Val Leu Gly Asn Leu Gln Leu Asn Leu Leu Ser  
 25 465 470 475 480  
 Lys Ser Lys Val Tyr Glu Asp Pro Ala Leu Ser Ala Ile Phe Leu His  
 485 490 495  
 Asn Asn Tyr Asn Tyr Ile Leu Lys Ser Leu Glu Lys Ser Glu Leu Ile  
 500 505 510  
 30 Gln Leu Val Ala Val Thr Gln Lys Thr Ala Glu Arg Ser Tyr Arg Glu  
 515 520 525  
 His Ile Glu Gln Gln Ile Gln Thr Tyr Gln Arg Ser Trp Leu Lys Val  
 530 535 540  
 Thr Asp Tyr Ile Ala Glu Lys Asn Leu Pro Val Phe Gln Pro Gly Val  
 35 545 550 555 560  
 Lys Leu Arg Asp Lys Glu Arg Gln Met Ile Lys Glu Arg Phe Lys Gly  
 565 570 575  
 Phe Asn Asp Gly Leu Glu Glu Leu Cys Lys Ile Gln Lys Ala Trp Ala  
 580 585 590  
 40 Ile Pro Asp Thr Glu Gln Arg Asp Lys Ile Arg Gln Ala Gln Lys Ser  
 595 600 605  
 Ile Val Lys Glu Thr Tyr Gly Ala Phe Leu His Arg Tyr Ser Ser Val  
 610 615 620  
 Pro Phe Thr Lys Asn Pro Glu Lys Tyr Ile Lys Tyr Arg Val Glu Gln  
 45 625 630 635 640

Val Gly Asp Met Ile Asp Arg Leu Phe Asp Thr Ser Ala  
 645 650

## (2) INFORMATION FOR SEQ ID NO:57:

5

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2496 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

10

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

15

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(C) INDIVIDUAL ISOLATE: p84 (rexo84)

20

## (ix) FEATURE:

(A) NAME/KEY: Coding Sequence

(B) LOCATION: 156...2303

25

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

GGCACGAGGC GCAATGCAGC GCTCGGCCCC GGCGCAGGCG TTTGGGTGTT GCTCCCGCCC  
 60

30 CGCTTCCTCT ACCGCTTCTC CACCGGGCTC GGGTGGGGCC GGGAGGCGGC GAGTGACAGC 120  
 CACTGCGGGC CACGGGAGCA GGGCGGGTGG CAGCC ATG TCG GAC AGC GGG GCG 173

Met Ser Asp Ser Gly Ala

1

5

35 AGC CGC CTG CGG AGG CAG CTG GAG TCG GGG GGC TTC GAG GCG CGG CTG 221  
 Ser Arg Leu Arg Arg Gln Leu Glu Ser Gly Gly Phe Glu Ala Arg Leu

10

15

20

40 TAC GTG AAG CAA CTG TCG CAG CAG TCG GAC GGC GAC CGC GAC CTG CAG 269  
 Tyr Val Lys Gln Leu Ser Gln Gln Ser Asp Gly Asp Arg Asp Leu Gln

25

30

35

GAG CAC CGA CAG CGG GTG CAG GCG CTG GCG GAG GAG ACG GCG CAG AAC 317  
 Glu His Arg Gln Arg Val Gln Ala Leu Ala Glu Glu Thr Ala Gln Asn

45

40

45

50

95

	CTG AAG CGC AAC GTC TAC CAG AAC TAC CGG CAG TTC ATC GAG ACG GCG	365
	Leu Lys Arg Asn Val Tyr Gln Asn Tyr Arg Gln Phe Ile Glu Thr Ala	
	55 60 65 70	
5	CGC GAG ATC TCC TAC CTG GAG AGC GAG ATG TAC CAG CTG AGC CAC CTG	413
	Arg Glu Ile Ser Tyr Leu Glu Ser Glu Met Tyr Gln Leu Ser His Leu	
	75 80 85	
10	CTG ACG GAG CAG AAG AGC AGT CTG GAG AGC ATC CCG CTG GCA CTG CTG	461
	Leu Thr Glu Gln Lys Ser Ser Leu Glu Ser Ile Pro Leu Ala Leu Leu	
	90 95 100	
15	CCC GCC GCT GCC GCG GGC GCA TCC ACG GGT GAG GAC ACG GCG GGC GCG	509
	Pro Ala Ala Ala Ala Gly Ala Ser Thr Gly Glu Asp Thr Ala Gly Ala	
	105 110 115	
20	GGA CCG CGG GAG CGC GGG GCG GCC CAG GCG GGC TTT CTT CCG GGG CCG	557
	Gly Pro Arg Glu Arg Gly Ala Ala Gln Ala Gly Phe Leu Pro Gly Pro	
	120 125 130	
25	GCC GGA GTC CCT CGC GAG GGC CCC GGG ACG GGA GAG GAG GGC AAG CAG	605
	Ala Gly Val Pro Arg Glu Gly Pro Gly Thr Gly Glu Glu Gly Lys Gln	
	135 140 145 150	
30	CGA ACG CTC ACC ACG CTG CTG GAG AAG GTG GAG GGC TGC AGG GAC CTG	653
	Arg Thr Leu Thr Thr Leu Leu Glu Lys Val Glu Gly Cys Arg Asp Leu	
	155 160 165	
35	CTG GAG ACA CCA GGT CAG TAC CTT GTG TAC AAC GGG GAC CTG GTG GAA	701
	Leu Glu Thr Pro Gly Gln Tyr Leu Val Tyr Asn Gly Asp Leu Val Glu	
	170 175 180	
40	TAC GAG GCA GAC CAT ATG GCC CAG CTG CAG CGA GTG CAT GGC TTC CTC	749
	Tyr Glu Ala Asp His Met Ala Gln Leu Gln Arg Val His Gly Phe Leu	
	185 190 195	
45	ATG AAC GAC TGT CTG CTG GTG GCC ACT TGG CTA CCA CAA CGG CGA GGC	797
	Met Asn Asp Cys Leu Leu Val Ala Thr Trp Leu Pro Gln Arg Arg Gly	
	200 205 210	
50	ATG TAC CGC TAC AAC GCA CTC TAC CCA CTG GAC CGC CTG GCT GTG GTC	845
	Met Tyr Arg Tyr Asn Ala Leu Tyr Pro Leu Asp Arg Leu Ala Val Val	
	215 220 225 230	
55	AAC GTC AAA GAC AAT CCA CCT ATG AAA GAC ATG TTC AAG CTG CTC ATG	893

96

	Asn Val Lys Asp Asn Pro Pro Met Lys Asp Met Phe Lys Leu Leu Met	
	235 240 245	
	TTC CCT GAG AGC CGA ATC TTT CAA GCG GAA AAT GCC AAG ATT AAA CGC	941
5	Phe Pro Glu Ser Arg Ile Phe Gln Ala Glu Asn Ala Lys Ile Lys Arg	
	250 255 260	
	GAG TGG CTG GAA GTG TTG GAG GAA ACC AAG AGG GCG CTC AGC GAC AAG	989
10	Glu Trp Leu Glu Val Leu Glu Glu Thr Lys Arg Ala Leu Ser Asp Lys	
	265 270 275	
	AGG AGA CGG GAG CAG GAG GAG GCA GCC GCC TTG CGT GCG CCA CCA CCG	1037
	Arg Arg Arg Glu Gln Glu Glu Ala Ala Ala Leu Arg Ala Pro Pro Pro	
15	280 285 290	
	GTC ACT TCC AAG GGC AGC AAC CCG TTT GAG GAT GAG GCC GAG GAG GAA	1085
	Val Thr Ser Lys Gly Ser Asn Pro Phe Glu Asp Glu Ala Glu Glu Glu	
	295 300 305 310	
20	CTG GCC ACC CCG GAG GCA GAG GAG GAA AAG GTT GAC CTT TCC ATG GAG	1133
	Leu Ala Thr Pro Glu Ala Glu Glu Glu Lys Val Asp Leu Ser Met Glu	
	315 320 325	
25	TGG ATC CAG GAG TTG CCG GAA GAC CTG GAT GTT TGT ATT GCG CAG AGG	1181
	Trp Ile Gln Glu Leu Pro Glu Asp Leu Asp Val Cys Ile Ala Gln Arg	
	330 335 340	
	GAC TTT GAG GGT GCC GTG GAC TTG CTG GAC AAA TTA AAT CAC TAT CTT	1229
30	Asp Phe Glu Gly Ala Val Asp Leu Leu Asp Lys Leu Asn His Tyr Leu	
	345 350 355	
	GAA GAT AAG CCC AGC CCA CCT TCT GTG AAA GAG CTG AGG GCC AAA GTG	1277
	Glu Asp Lys Pro Ser Pro Pro Ser Val Lys Glu Leu Arg Ala Lys Val	
35	360 365 370	
	GAT GAA CGA GTG CGA CAG CTC ACC GAG GTG CTT GTG TTT GAG CTC TCC	1325
	Asp Glu Arg Val Arg Gln Leu Thr Glu Val Leu Val Phe Glu Leu Ser	
	375 380 385 390	
40	CCG GAT CGC TCT CTG AGA GGT GGC CCT AAG GCT ACT CGA AGG GCA GTG	1373
	Pro Asp Arg Ser Leu Arg Gly Gly Pro Lys Ala Thr Arg Arg Ala Val	
	395 400 405	
45	TCT CAA CTG ATC CGT CTC GGC CAG TGC ACT AAG GCT TGT GAG CTG TTT	1421

97

	Ser Gln Leu Ile Arg Leu Gly Gln Cys Thr Lys Ala Cys Glu Leu Phe	
	410 415 420	
5	CTG AGG AAC AGG GCG GCA GCT GTG CAT ACT GCC ATC CGC CAG CTT CGA	1469
	Leu Arg Asn Arg Ala Ala Ala Val His Thr Ala Ile Arg Gln Leu Arg	
	425 430 435	
10	ATT GAG GGC GCC ACT CTG CTC TAT ATT CAC AAG CTG TGC CAT GTC TTC	1517
	Ile Glu Gly Ala Thr Leu Leu Tyr Ile His Lys Leu Cys His Val Phe	
	440 445 450	
15	TTT ACC AGC CTC CTA GAG ACT GCA CGG GAG TTT GAG ACA GAC TTT GCA	1565
	Phe Thr Ser Leu Leu Glu Thr Ala Arg Glu Phe Glu Thr Asp Phe Ala	
	455 460 465 470	
20	GGC ACG GAC AGT GGC TGC TAC TCT GCC TTT GTG GTC TGG GCA AGG TCT	1613
	Gly Thr Asp Ser Gly Cys Tyr Ser Ala Phe Val Val Trp Ala Arg Ser	
	475 480 485	
	GCC ATG GGC ATG TTC GTG GAT GCT TTC AGC AAG CAG GTT TTT GAC AGC	1661
	Ala Met Gly Met Phe Val Asp Ala Phe Ser Lys Gln Val Phe Asp Ser	
	490 495 500	
25	AAG GAG AGC CTG TCC ACT GCT GCC GAG TGT GTG AAG GTA GCC AAG GAG	1709
	Lys Glu Ser Leu Ser Thr Ala Ala Glu Cys Val Lys Val Ala Lys Glu	
	505 510 515	
30	CAC TGC CAG CAG CTG GGA GAG ATT GGG CTG GAT CTC ACC TTC ATC ATC	1757
	His Cys Gln Gln Leu Gly Glu Ile Gly Leu Asp Leu Thr Phe Ile Ile	
	520 525 530	
35	CAT GCC CTC CTG GTG AAG GAC ATC CAG GGG GCC TTG CTC AGT TAC AAG	1805
	His Ala Leu Leu Val Lys Asp Ile Gln Gly Ala Leu Leu Ser Tyr Lys	
	535 540 545 550	
40	GAG ATT ATC ATT GAA GCC ACC AAG CAC CGA AAC TCG GAG GAG ATG TGG	1853
	Glu Ile Ile Ile Glu Ala Thr Lys His Arg Asn Ser Glu Glu Met Trp	
	555 560 565	
	CGT CGG ATG AAC CTG ATG ACT CCT GAG GCC CTG GGC AAG CTC AAG GAG	1901
	Arg Arg Met Asn Leu Met Thr Pro Glu Ala Leu Gly Lys Leu Lys Glu	
	570 575 580	
45	GAG ATG AGA AGT TGC GGG GTC AGT AAC TTT GAG CAG TAT ACG GGG GAT	1949

Glu Met Arg Ser Cys Gly Val Ser Asn Phe Glu Gln Tyr Thr Gly Asp  
585 590 595

5 GAC TGC TGG GTG AAC CTG AGC TAC ACT GTG GTA GCC TTC ACC AAG CAG 1997  
Asp Cys Trp Val Asn Leu Ser Tyr Thr Val Val Ala Phe Thr Lys Gln  
600 605 610

10 ACC ATG GGC TTC CTG GAG GAG GCA CTG AAA CTG TAC TTC CCA GAG CTA 2045  
Thr Met Gly Phe Leu Glu Glu Ala Leu Lys Leu Tyr Phe Pro Glu Leu  
615 620 625 630

15 CAC ATG GTG CTC CTG GAG AGT CTG GTG GAA GTC ATA CTG GTT GCT GTG 2093  
His Met Val Leu Leu Glu Ser Leu Val Glu Val Ile Leu Val Ala Val  
635 640 645

20 CAG CAT GTG GAC TAC AGC CTG CGC TGT GAG CAG GAC CCA GAG AAG AAG 2141  
Gln His Val Asp Tyr Ser Leu Arg Cys Glu Gln Asp Pro Glu Lys Lys  
650 655 660

ACT TTC ATC CGG CAG AAC GCA TCC TTC CTT TAC GAC ACT GTC CTC CCA 2189  
Thr Phe Ile Arg Gln Asn Ala Ser Phe Leu Tyr Asp Thr Val Leu Pro  
665 670 675

25 GTG GTG GAG AGG AGG TTT GAG GAA GGT GTA GGG AAG CCT GCC AAG CAG 2237  
Val Val Glu Arg Arg Phe Glu Glu Gly Val Gly Lys Pro Ala Lys Gln  
680 685 690

30 CTG CAG GAC CTT CGG AAT GCA TCG AGA CTC CTG CGG GTG AAT CCT GAG 2285  
Leu Gln Asp Leu Arg Asn Ala Ser Arg Leu Leu Arg Val Asn Pro Glu  
695 700 705 710

35 AGC ACC ACG TCT GTG GTC TGAGGCTCTG GTTCGATGCT GTGTGTGCGT CTGTCTGT 2341  
Ser Thr Thr Ser Val Val  
715

40 CTGTCTGTCT GTATGTATGT ATGTATGTAT GTATGTCTGT ACACACACAT GTACTACAGA 2401  
TGCATTGTTT GTATTCACAT TACACAATAT CAACACATTC CTCAGACACA AACAGCATAT 2461  
AAAAGTGCCC TCTTAAAAAT GCAGTCAAAA AAAAA 2496

(2) INFORMATION FOR SEQ ID NO:58:

(i) SEQUENCE CHARACTERISTICS:

45 (A) LENGTH: 716 amino acids



99

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

```

Met Ser Asp Ser Gly Ala Ser Arg Leu Arg Arg Gln Leu Glu Ser Gly
 1           5           10           15
10 Gly Phe Glu Ala Arg Leu Tyr Val Lys Gln Leu Ser Gln Gln Ser Asp
    20           25           30
Gly Asp Arg Asp Leu Gln Glu His Arg Gln Arg Val Gln Ala Leu Ala
    35           40           45
Glu Glu Thr Ala Gln Asn Leu Lys Arg Asn Val Tyr Gln Asn Tyr Arg
15    50           55           60
Gln Phe Ile Glu Thr Ala Arg Glu Ile Ser Tyr Leu Glu Ser Glu Met
65    70           75           80
Tyr Gln Leu Ser His Leu Leu Thr Glu Gln Lys Ser Ser Leu Glu Ser
    85           90           95
20 Ile Pro Leu Ala Leu Leu Pro Ala Ala Ala Gly Ala Ser Thr Gly
    100          105          110
Glu Asp Thr Ala Gly Ala Gly Pro Arg Glu Arg Gly Ala Ala Gln Ala
    115          120          125
Gly Phe Leu Pro Gly Pro Ala Gly Val Pro Arg Glu Gly Pro Gly Thr
25    130          135          140
Gly Glu Glu Gly Lys Gln Arg Thr Leu Thr Thr Leu Leu Glu Lys Val
145          150          155          160
Glu Gly Cys Arg Asp Leu Leu Glu Thr Pro Gly Gln Tyr Leu Val Tyr
    165          170          175
30 Asn Gly Asp Leu Val Glu Tyr Glu Ala Asp His Met Ala Gln Leu Gln
    180          185          190
Arg Val His Gly Phe Leu Met Asn Asp Cys Leu Leu Val Ala Thr Trp
    195          200          205
Leu Pro Gln Arg Arg Gly Met Tyr Arg Tyr Asn Ala Leu Tyr Pro Leu
35    210          215          220
Asp Arg Leu Ala Val Val Asn Val Lys Asp Asn Pro Pro Met Lys Asp
225          230          235          240
Met Phe Lys Leu Leu Met Phe Pro Glu Ser Arg Ile Phe Gln Ala Glu
    245          250          255
40 Asn Ala Lys Ile Lys Arg Glu Trp Leu Glu Val Leu Glu Glu Thr Lys
    260          265          270
Arg Ala Leu Ser Asp Lys Arg Arg Arg Glu Gln Glu Glu Ala Ala Ala
    275          280          285
Leu Arg Ala Pro Pro Pro Val Thr Ser Lys Gly Ser Asn Pro Phe Glu
45    290          295          300

```

100

Asp Glu Ala Glu Glu Glu Leu Ala Thr Pro Glu Ala Glu Glu Glu Lys  
 305 310 315 320  
 Val Asp Leu Ser Met Glu Trp Ile Gln Glu Leu Pro Glu Asp Leu Asp  
 325 330 335  
 5 Val Cys Ile Ala Gln Arg Asp Phe Glu Gly Ala Val Asp Leu Leu Asp  
 340 345 350  
 Lys Leu Asn His Tyr Leu Glu Asp Lys Pro Ser Pro Pro Ser Val Lys  
 355 360 365  
 Glu Leu Arg Ala Lys Val Asp Glu Arg Val Arg Gln Leu Thr Glu Val  
 10 370 375 380  
 Leu Val Phe Glu Leu Ser Pro Asp Arg Ser Leu Arg Gly Gly Pro Lys  
 385 390 395 400  
 Ala Thr Arg Arg Ala Val Ser Gln Leu Ile Arg Leu Gly Gln Cys Thr  
 405 410 415  
 15 Lys Ala Cys Glu Leu Phe Leu Arg Asn Arg Ala Ala Ala Val His Thr  
 420 425 430  
 Ala Ile Arg Gln Leu Arg Ile Glu Gly Ala Thr Leu Leu Tyr Ile His  
 435 440 445  
 Lys Leu Cys His Val Phe Phe Thr Ser Leu Leu Glu Thr Ala Arg Glu  
 20 450 455 460  
 Phe Glu Thr Asp Phe Ala Gly Thr Asp Ser Gly Cys Tyr Ser Ala Phe  
 465 470 475 480  
 Val Val Trp Ala Arg Ser Ala Met Gly Met Phe Val Asp Ala Phe Ser  
 485 490 495  
 25 Lys Gln Val Phe Asp Ser Lys Glu Ser Leu Ser Thr Ala Ala Glu Cys  
 500 505 510  
  
 Val Lys Val Ala Lys Glu His Cys Gln Gln Leu Gly Glu Ile Gly Leu  
 515 520 525  
 30 Asp Leu Thr Phe Ile Ile His Ala Leu Leu Val Lys Asp Ile Gln Gly  
 530 535 540  
 Ala Leu Leu Ser Tyr Lys Glu Ile Ile Ile Glu Ala Thr Lys His Arg  
 545 550 555 560  
 Asn Ser Glu Glu Met Trp Arg Arg Met Asn Leu Met Thr Pro Glu Ala  
 565 570 575  
 35 Leu Gly Lys Leu Lys Glu Glu Met Arg Ser Cys Gly Val Ser Asn Phe  
 580 585 590  
 Glu Gln Tyr Thr Gly Asp Asp Cys Trp Val Asn Leu Ser Tyr Thr Val  
 595 600 605  
 40 Val Ala Phe Thr Lys Gln Thr Met Gly Phe Leu Glu Glu Ala Leu Lys  
 610 615 620  
 Leu Tyr Phe Pro Glu Leu His Met Val Leu Leu Glu Ser Leu Val Glu  
 625 630 635 640  
 Val Ile Leu Val Ala Val Gln His Val Asp Tyr Ser Leu Arg Cys Glu  
 45 645 650 655

101

Gln Asp Pro Glu Lys Lys Thr Phe Ile Arg Gln Asn Ala Ser Phe Leu  
                   660                  665                  670  
 Tyr Asp Thr Val Leu Pro Val Val Glu Arg Arg Phe Glu Glu Gly Val  
                   675                  680                  685  
 5 Gly Lys Pro Ala Lys Gln Leu Gln Asp Leu Arg Asn Ala Ser Arg Leu  
                   690                  695                  700  
 Leu Arg Val Asn Pro Glu Ser Thr Thr Ser Val Val  
                   705                  710                  715

10

(2) INFORMATION FOR SEQ ID NO:59:

(i) SEQUENCE CHARACTERISTICS:

- 15 (A) LENGTH: 3059 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

20

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

25

(vi) ORIGINAL SOURCE:

(C) INDIVIDUAL ISOLATE: p96 (rsec15)

(ix) FEATURE:

- 30 (A) NAME/KEY: Coding Sequence  
 (B) LOCATION: 341...2806

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

35 CCGCGTCGAC CCGGCACAAA CACACCATGC ACCCTGACAG TAAATTTATA AAAAATTTTT  
 60  
 TTCAACTATA GCACAAAAAA TGTGTATCAC TTAACAGGGA TGACAAATGG CTGTCACTTC 120  
 AGTCACTATA ATGCACAATA TATATATGCC TTCTACTTAT GAAGCACGTT TACTGGAAAA 180  
 CAACTGTTAT GCAGACGGCA TCCTCATTTT TCTTATGTTT GGTCTCTTGT CCCTTGAGAG 240  
 40 ATCATGTCCA GTAAGGGATC TACAACAGTT CCTTGTGTGT GTTAACAATG ATGAAATCAC 300  
 AATTGATGTG TTTTCTGAG TCTGTCCCA CAGTTAAACA ATG CCT TTC TCC CCG 355  
   Met Pro Phe Ser Pro  
   1                  5  
 45 GGG CGC CGC GCC CGG CCG GCG CCT CCG CTT CCA GCC AAG ATG GCG GAG 403

102

	Gly Arg Arg Ala Arg Pro Ala Pro Pro Leu Pro Ala Lys Met Ala Glu	
	10 15 20	
5	AGC GGC GAG GCT CTG GGT ACG GTC CCG GAG CAC GAG CGG ATC TTG CAG Ser Gly Glu Ala Leu Gly Thr Val Pro Glu His Glu Arg Ile Leu Gln	451
	25 30 35	
10	GAG ATC GAA AGC ACG GAC ACC GCC TGC GTA GGG CCC ACC CTT CGG TCT Glu Ile Glu Ser Thr Asp Thr Ala Cys Val Gly Pro Thr Leu Arg Ser	499
	40 45 50	
15	GTG TAT GAT GGT CAG CCA AAT GCA CAC AAG AAG TTC ATG GAA AAG CTA Val Tyr Asp Gly Gln Pro Asn Ala His Lys Lys Phe Met Glu Lys Leu	547
	55 60 65	
20	GAT GCC TGC ATC CGC AAC CAT GAC AAG GAG ATT GAG AAA ATG TGT AAC Asp Ala Cys Ile Arg Asn His Asp Lys Glu Ile Glu Lys Met Cys Asn	595
	70 75 80 85	
25	TTC CAT CAC CAG GGC TTT GTA GAT GCT ATC ACA GAG CTC CTC AAA GTA Phe His His Gln Gly Phe Val Asp Ala Ile Thr Glu Leu Leu Lys Val	643
	90 95 100	
30	AGG GCG GAC GCA GAA AAG CTG AAG GTC CAG GTT ACT GAT ACC AAC CGA Arg Ala Asp Ala Glu Lys Leu Lys Val Gln Val Thr Asp Thr Asn Arg	691
	105 110 115	
35	AGG TTC CAA GAT GCT GGA AAA GAG GTG ATA GAA CAA ACA GAA GAT ATT Arg Phe Gln Asp Ala Gly Lys Glu Val Ile Glu Gln Thr Glu Asp Ile	739
	120 125 130	
40	ATT CGG TGT AGA ATT CAG CAG AGA AAT ATT ACA ACC GTA GTA GAG AAA Ile Arg Cys Arg Ile Gln Gln Arg Asn Ile Thr Thr Val Val Glu Lys	787
	135 140 145	
45	TTG CAG TTA TGC CTG CCA GTG TTG GAG ATG TAC AGT AAA CTG AAG GAA Leu Gln Leu Cys Leu Pro Val Leu Glu Met Tyr Ser Lys Leu Lys Glu	835
	150 155 160 165	
50	CAG ATG AGC ATG CAG AGG TAT TAT TCT GCA CTT AAA ACA ATG GAA CAG Gln Met Ser Met Gln Arg Tyr Tyr Ser Ala Leu Lys Thr Met Glu Gln	883
	170 175 180	
55	TTA GAG AAC GTA TAC TTT CCT CGG GTT AGC CAG TAC CGC TTC TGT CAG Leu Glu Asn Val Tyr Phe Pro Arg Val Ser Gln Tyr Arg Phe Cys Gln	931

		103	
	185	190	195
	CTC ATG ATG GAC ACA CTT CCT AAA CTC CGG GAG GAT ATC AAA GAC ATC		979
	Leu Met Met Asp Thr Leu Pro Lys Leu Arg Glu Asp Ile Lys Asp Ile		
5	200	205	210
	TCC ATG TCT GAC CTC AAG GAC TTT CTG GAA AGC ATT CGA AAG CAT TCT		1027
	Ser Met Ser Asp Leu Lys Asp Phe Leu Glu Ser Ile Arg Lys His Ser		
	215	220	225
10	GAC AAA ATA GGG GAA ACA GCC ATG AAA CAG GCA CAA CAG CAG AAG AGC		1075
	Asp Lys Ile Gly Glu Thr Ala Met Lys Gln Ala Gln Gln Gln Lys Ser		
	230	235	240 245
	TTC AGT ATT GCT GTA CAG AAG CAA ACC AAT ATG AGG TTT GGG AAA AAT		1123
	Phe Ser Ile Ala Val Gln Lys Gln Thr Asn Met Arg Phe Gly Lys Asn		
	250	255	260
20	ATG CAT GTA AAC AAT GAC AGG ACT CTA GAG GAA AAG AGT GAC ATC ATA		1171
	Met His Val Asn Asn Asp Arg Thr Leu Glu Glu Lys Ser Asp Ile Ile		
	265	270	275
	CTG AAG CAC ACA CTT GAG GAG GAG GCC GAG AAT GAC GAG GAG GTC TTA		1219
25	Leu Lys His Thr Leu Glu Glu Glu Ala Glu Asn Asp Glu Glu Val Leu		
	280	285	290
	ACT GTT CAA GAT CTC GTT GAC TTT TCC CCT GTT TAC CGA TGC TCA CAC		1267
	Thr Val Gln Asp Leu Val Asp Phe Ser Pro Val Tyr Arg Cys Ser His		
30	295	300	305
	ATT TAT TCT GCT TTG GGT GAT GAA GAA ACA TTT GAA AAT TAT TAC CGG		1315
	Ile Tyr Ser Ala Leu Gly Asp Glu Glu Thr Phe Glu Asn Tyr Tyr Arg		
	310	315	320 325
35	AAG CAA AGG AAG AAG CAG GCG CGC CTG GTT CTG CAG CCG CAG TCG AGT		1363
	Lys Gln Arg Lys Lys Gln Ala Arg Leu Val Leu Gln Pro Gln Ser Ser		
	330	335	340
40	GTG CAT GAA ACA GTT GAT GGC TAT AGA AGA TAT TTC ACT CAG ATT GTA		1411
	Val His Glu Thr Val Asp Gly Tyr Arg Arg Tyr Phe Thr Gln Ile Val		
	345	350	355
	GGG TTC TTT GTT GTG GAA GAT CAC ATT TTA CAT GTG ACC CAA GGA TTA		1459
45	Gly Phe Phe Val Val Glu Asp His Ile Leu His Val Thr Gln Gly Leu		

105

	535	540	545	
	TGC TTA CTG AAC CTC ATT AGA AAA CCT CAC ATA GGT TTG ACA GAG CTG	2035		
	Cys Leu Leu Asn Leu Ile Arg Lys Pro His Ile Gly Leu Thr Glu Leu			
5	550	555	560	565
	GTG CAA ATT ATC ATA AAC ACA ACA CAC CTG GAG CAG GCT TGC AAA TAC	2083		
	Val Gln Ile Ile Ile Asn Thr Thr His Leu Glu Gln Ala Cys Lys Tyr			
10	570	575	580	
	CTG GAG GAC TTT ATA ACT AAC ATT ACA AAT ATT TCT CAA GAA ACC GTT	2131		
	Leu Glu Asp Phe Ile Thr Asn Ile Thr Asn Ile Ser Gln Glu Thr Val			
	585	590	595	
15				
	CAC ACC ACA AGA CTT TAT GGA CTT TCT ACT TTT AAG GAT GCA CGA CAT	2179		
	His Thr Thr Arg Leu Tyr Gly Leu Ser Thr Phe Lys Asp Ala Arg His			
	600	605	610	
20				
	GCA GCG GAA GGA GAA ATA TAT ACC AAA CTC AAT CAG AAG ATT GAT GAG	2227		
	Ala Ala Glu Gly Glu Ile Tyr Thr Lys Leu Asn Gln Lys Ile Asp Glu			
	615	620	625	
25				
	TTT GTC CAG CTG GCC GAC TAC GAC TGG ACC ATG GCT GAG TCG GAT GGG	2275		
	Phe Val Gln Leu Ala Asp Tyr Asp Trp Thr Met Ala Glu Ser Asp Gly			
	630	635	640	645
30				
	AGA GCA AGT GGT TAT TTA ATG GAT CTT ATT AAT TTT TTG AGA AGC ATC	2323		
	Arg Ala Ser Gly Tyr Leu Met Asp Leu Ile Asn Phe Leu Arg Ser Ile			
	650	655	660	
35				
	TTT CAA GTA TTT ACC CAC TTG CCT GGG AAA GTT GCC CAG ACA GCT TGC	2371		
	Phe Gln Val Phe Thr His Leu Pro Gly Lys Val Ala Gln Thr Ala Cys			
	665	670	675	
40				
	ATG TCA GCC TGC CAA CAT CTG TCC ACA TCC TTG ATG CAG ATG CTC CTG	2419		
	Met Ser Ala Cys Gln His Leu Ser Thr Ser Leu Met Gln Met Leu Leu			
	680	685	690	
45				
	GAC AGC GAG TTA AAG CAG ATC AGC ATG GGG GCC GTA CAG CAG TTT AAC	2467		
	Asp Ser Glu Leu Lys Gln Ile Ser Met Gly Ala Val Gln Gln Phe Asn			
	695	700	705	
	CTA GAC GTC ATA CAG TGC GAA TTG TTT GCC AGC TCT GAG CCT GTG CCA	2515		
	Leu Asp Val Ile Gln Cys Glu Leu Phe Ala Ser Ser Glu Pro Val Pro			

106

	710	715	720	725		
	GGA TTC CAA GGG GAC ACC CTG CAG CTG GCC TTC ATT GAC CTC AGA CAG				2563	
5	Gly Phe Gln Gly Asp Thr Leu Gln Leu Ala Phe Ile Asp Leu Arg Gln					
		730	735	740		
	CTC CTT GAC TTG TTT ATG GTT TGG GAC TGG TCT ACT TAC CTG GCT GAC				2611	
	Leu Leu Asp Leu Phe Met Val Trp Asp Trp Ser Thr Tyr Leu Ala Asp					
10		745	750	755		
	TAT GGG CAG CCA GCT TCC AAG TAC CTC CGG GTG AAC CCA CAC GCC GCC				2659	
	Tyr Gly Gln Pro Ala Ser Lys Tyr Leu Arg Val Asn Pro His Ala Ala					
		760	765	770		
15	CTT ACG CTT TTG GAG AAG ATG AAG GAC ACC AGT AAG AAG AAT AAC ATA					2707
	Leu Thr Leu Leu Glu Lys Met Lys Asp Thr Ser Lys Lys Asn Asn Ile					
		775	780	785		
20	TTT GCC CAG TTC CGG AAG AAC GAC CGG GAC AGA CAG AAG CTG ATC GAG				2755	
	Phe Ala Gln Phe Arg Lys Asn Asp Arg Asp Arg Gln Lys Leu Ile Glu					
		790	795	800	805	
	ACA GTG GTG AAA CAG CTC AGA GGC TTG GTG ACC GGT ATG TCC CAG CAC				2803	
25	Thr Val Val Lys Gln Leu Arg Gly Leu Val Thr Gly Met Ser Gln His					
		810	815	820		
	ATG TAGACCAGAT CCATGCCCGA GTGCTTCTCT GTGGCCAGCA CCGGTCATGA GACAAT				2862	
	Met					
30	AATTTGTTTA CAGAAGCCAA AACTGTAGT AGAGGAAAGA TACTAATGAG GGCTTAAACA					2922
	ATAAGGACAG CAAACATCAT TTGTATGGAC TTTTAATAAC TGAATACTAT TTTATATATA					2982
	AAAATGTAAC ATTAATTTTT TCCAATTTTA GAAGAAAAGT CAAAAGTGTA ATATATAAAC					3042
	CCAATAAATT GTGTATG					3059

35

## (2) INFORMATION FOR SEQ ID NO:60:

## (i) SEQUENCE CHARACTERISTICS:

40

- (A) LENGTH: 822 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

45

107

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

```

Met Pro Phe Ser Pro Gly Arg Arg Ala Arg Pro Ala Pro Pro Leu Pro
  1           5           10           15
5  Ala Lys Met Ala Glu Ser Gly Glu Ala Leu Gly Thr Val Pro Glu His
    20           25           30
Glu Arg Ile Leu Gln Glu Ile Glu Ser Thr Asp Thr Ala Cys Val Gly
    35           40           45
Pro Thr Leu Arg Ser Val Tyr Asp Gly Gln Pro Asn Ala His Lys Lys
10   50           55           60
Phe Met Glu Lys Leu Asp Ala Cys Ile Arg Asn His Asp Lys Glu Ile
65           70           75           80
Glu Lys Met Cys Asn Phe His His Gln Gly Phe Val Asp Ala Ile Thr
    85           90           95
15  Glu Leu Leu Lys Val Arg Ala Asp Ala Glu Lys Leu Lys Val Gln Val
    100          105          110
Thr Asp Thr Asn Arg Arg Phe Gln Asp Ala Gly Lys Glu Val Ile Glu
    115          120          125
Gln Thr Glu Asp Ile Ile Arg Cys Arg Ile Gln Gln Arg Asn Ile Thr
20   130          135          140
Thr Val Val Glu Lys Leu Gln Leu Cys Leu Pro Val Leu Glu Met Tyr
145          150          155          160
Ser Lys Leu Lys Glu Gln Met Ser Met Gln Arg Tyr Tyr Ser Ala Leu
    165          170          175
25  Lys Thr Met Glu Gln Leu Glu Asn Val Tyr Phe Pro Arg Val Ser Gln
    180          185          190
Tyr Arg Phe Cys Gln Leu Met Met Asp Thr Leu Pro Lys Leu Arg Glu
    195          200          205
Asp Ile Lys Asp Ile Ser Met Ser Asp Leu Lys Asp Phe Leu Glu Ser
30   210          215          220
Ile Arg Lys His Ser Asp Lys Ile Gly Glu Thr Ala Met Lys Gln Ala
225          230          235          240

Gln Gln Gln Lys Ser Phe Ser Ile Ala Val Gln Lys Gln Thr Asn Met
35   245          250          255
Arg Phe Gly Lys Asn Met His Val Asn Asn Asp Arg Thr Leu Glu Glu
    260          265          270
Lys Ser Asp Ile Ile Leu Lys His Thr Leu Glu Glu Glu Ala Glu Asn
    275          280          285
40  Asp Glu Glu Val Leu Thr Val Gln Asp Leu Val Asp Phe Ser Pro Val
    290          295          300
Tyr Arg Cys Ser His Ile Tyr Ser Ala Leu Gly Asp Glu Glu Thr Phe
305          310          315          320
Glu Asn Tyr Tyr Arg Lys Gln Arg Lys Lys Gln Ala Arg Leu Val Leu
45   325          330          335

```



108

Gln Pro Gln Ser Ser Val His Glu Thr Val Asp Gly Tyr Arg Arg Tyr  
                   340                  345                  350  
 Phe Thr Gln Ile Val Gly Phe Phe Val Val Glu Asp His Ile Leu His  
                   355                  360                  365  
 5 Val Thr Gln Gly Leu Val Thr Arg Ala Tyr Thr Asp Glu Leu Trp Asn  
                   370                  375                  380  
 Met Ala Leu Ser Lys Ile Ile Ala Val Leu Arg Ala His Ser Ser Tyr  
                   385                  390                  395                  400  
 Cys Thr Asp Pro Asp Leu Val Leu Glu Leu Lys Xaa Leu Ile Val Ile  
 10                                   405                  410                  415  
 Phe Ala Asp Thr Leu Gln Gly Tyr Gly Phe Ser Val Asn Arg Leu Phe  
                   420                  425                  430  
 Asp Leu Leu Phe Glu Ile Arg Asp Gln Tyr Asn Glu Thr Leu Leu Lys  
                   435                  440                  445  
 15 Lys Trp Ala Gly Ile Phe Arg Asp Ile Phe Glu Glu Asp Asn Tyr Ser  
                   450                  455                  460  
 Pro Ile Pro Ile Gly Ser Glu Glu Glu Tyr Lys Met Val Ile Ser Lys  
                   465                  470                  475                  480  
 Phe Pro Phe Gln Asp Pro Asp Leu Glu Lys Gln Ser Phe Pro Lys Lys  
 20                                   485                  490                  495  
 Phe Pro Met Ser Gln Ser Val Pro Leu Ile Tyr Ile Gln Val Lys Glu  
                   500                  505                  510  
 Phe Ile Tyr Ala Ser Leu Lys Phe Ser Glu Ser Leu His Arg Ser Ser  
                   515                  520                  525  
 25 Thr Glu Ile Asp Asp Met Leu Arg Lys Ser Thr Asn Leu Leu Leu Thr  
                   530                  535                  540  
 Arg Ile Leu Ser Ser Cys Leu Leu Asn Leu Ile Arg Lys Pro His Ile  
                   545                  550                  555                  560  
  
 30 Gly Leu Thr Glu Leu Val Gln Ile Ile Ile Asn Thr Thr His Leu Glu  
                   565                  570                  575  
 Gln Ala Cys Lys Tyr Leu Glu Asp Phe Ile Thr Asn Ile Thr Asn Ile  
                   580                  585                  590  
 Ser Gln Glu Thr Val His Thr Thr Arg Leu Tyr Gly Leu Ser Thr Phe  
 35                                   595                  600                  605  
 Lys Asp Ala Arg His Ala Ala Glu Gly Glu Ile Tyr Thr Lys Leu Asn  
                   610                  615                  620  
 Gln Lys Ile Asp Glu Phe Val Gln Leu Ala Asp Tyr Asp Trp Thr Met  
                   625                  630                  635                  640  
 40 Ala Glu Ser Asp Gly Arg Ala Ser Gly Tyr Leu Met Asp Leu Ile Asn  
                   645                  650                  655  
 Phe Leu Arg Ser Ile Phe Gln Val Phe Thr His Leu Pro Gly Lys Val  
                   660                  665                  670  
 Ala Gln Thr Ala Cys Met Ser Ala Cys Gln His Leu Ser Thr Ser Leu  
 45                                   675                  680                  685

109

Met Gln Met Leu Leu Asp Ser Glu Leu Lys Gln Ile Ser Met Gly Ala  
690 695 700  
Val Gln Gln Phe Asn Leu Asp Val Ile Gln Cys Glu Leu Phe Ala Ser  
705 710 715 720  
5 Ser Glu Pro Val Pro Gly Phe Gln Gly Asp Thr Leu Gln Leu Ala Phe  
725 730 735  
Ile Asp Leu Arg Gln Leu Leu Asp Leu Phe Met Val Trp Asp Trp Ser  
740 745 750  
Thr Tyr Leu Ala Asp Tyr Gly Gln Pro Ala Ser Lys Tyr Leu Arg Val  
10 755 760 765  
Asn Pro His Ala Ala Leu Thr Leu Leu Glu Lys Met Lys Asp Thr Ser  
770 775 780  
Lys Lys Asn Asn Ile Phe Ala Gln Phe Arg Lys Asn Asp Arg Asp Arg  
785 790 795 800  
15 Gln Lys Leu Ile Glu Thr Val Val Lys Gln Leu Arg Gly Leu Val Thr  
805 810 815  
Gly Met Ser Gln His Met  
820

20

(2) INFORMATION FOR SEQ ID NO:61:

(i) SEQUENCE CHARACTERISTICS:

- 25 (A) LENGTH: 4358 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

30

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

35

(vi) ORIGINAL SOURCE:

(C) INDIVIDUAL ISOLATE: p102 (rsec5)

(ix) FEATURE:

- 40 (A) NAME/KEY: Coding Sequence  
(B) LOCATION: 200...2971

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

45

110

ACTGCTGCGG CCCGAGCGG GTTGGCCCGG ACCTGACGCC GCCTGCAAGC TGCAGGACAG  
 60  
 AGGGCTGGAG ATGGCAGATC CCAGGAGCTA GCAGCCAGCC CAGCCGAATG GTGTCTCCAC 120  
 CTGGAGTGCA GGAGAGAGAG ACTGAAAACC AAAAAGGTTG CTTTATAAC TGGAACAGGA 180  
 5 GTTTGCTGCT CAAAGTACA ATG TCT CGA TCC CGG CAG CCT CCA CTT GTG ACA 232  
 Met Ser Arg Ser Arg Gln Pro Pro Leu Val Thr  
 1 5 10  
 GGC ATC TCT CCA AAT GAA GGG ATT CCT TGG ACC AAA GTC ACA ATT AGA 280  
 10 Gly Ile Ser Pro Asn Glu Gly Ile Pro Trp Thr Lys Val Thr Ile Arg  
 15 20 25  
 GGG GAG AAT CTG GGT ACT GGT CCC ACT GAC CTT ATA GGC TTG ACA ATT 328  
 Gly Glu Asn Leu Gly Thr Gly Pro Thr Asp Leu Ile Gly Leu Thr Ile  
 15 30 35 40  
 TGT GGA CAT AAT TGC CTC CTC ACG GCA GAA TGG ATG TCT GCA AGT AAA 376  
 Cys Gly His Asn Cys Leu Leu Thr Ala Glu Trp Met Ser Ala Ser Lys  
 45 50 55  
 20 ATT GTC TGT CGA GTG GGA CAA GCC AAA AAT GAC AAA GGA GAC ATT ATT 424  
 Ile Val Cys Arg Val Gly Gln Ala Lys Asn Asp Lys Gly Asp Ile Ile  
 60 65 70 75  
 25 GTC ACA ACC AAG TCA GGG GGC AGA GGA ACC TCA ACT GTC TCT TTT AAG 472  
 Val Thr Thr Lys Ser Gly Gly Arg Gly Thr Ser Thr Val Ser Phe Lys  
 80 85 90  
 CTA CTC AAA CCT GAA AAA ATA GGC ATT TTG GAT CAG TCT GCA GTG TGG 520  
 30 Leu Leu Lys Pro Glu Lys Ile Gly Ile Leu Asp Gln Ser Ala Val Trp  
 95 100 105  
 GTG GAT GAG ATG AAT TAT TAT GAT ATG CGG ACT GAC AGG AAT AAG GGG 568  
 Val Asp Glu Met Asn Tyr Tyr Asp Met Arg Thr Asp Arg Asn Lys Gly  
 35 110 115 120  
 ATT CCC CCC TTG TCA CTT CGT CCT GCG AAT CCT CTC GGC ATT GAA ATT 616  
 Ile Pro Pro Leu Ser Leu Arg Pro Ala Asn Pro Leu Gly Ile Glu Ile  
 125 130 135  
 40 GAG AAG TGT AAA CTT CCT CAG AAG AAC TTA GAA GTG TTA TTC CAT GGC 664  
 Glu Lys Cys Lys Leu Pro Gln Lys Asn Leu Glu Val Leu Phe His Gly  
 140 145 150 155  
 45 ATG AGT GCT GAT TTT ACG AGT GAG AAT TTT TCA GCA GCC TGG TAT CTC 712

111

	Met Ser Ala Asp Phe Thr Ser Glu Asn Phe Ser Ala Ala Trp Tyr Leu	
	160 165 170	
5	ATA GAG AAC CAC TCC AAC ACC AGT TTT GAG CAG CTC AAG ATG GCA GTT Ile Glu Asn His Ser Asn Thr Ser Phe Glu Gln Leu Lys Met Ala Val	760
	175 180 185	
10	ACC AAC CTC AAA AGG CAG GCC AAC AAG AAG AGC GAG GGT AGT CTA GCG Thr Asn Leu Lys Arg Gln Ala Asn Lys Lys Ser Glu Gly Ser Leu Ala	808
	190 195 200	
15	TAC GTG AAG GGT GGC CTC AGT ACC TTC TTT GAA GCT CAA GAT GCT CTC Tyr Val Lys Gly Gly Leu Ser Thr Phe Phe Glu Ala Gln Asp Ala Leu	856
	205 210 215	
	TCA GCT ATC CAT CAA AAA CTA GAA GCA GAC GGA ACA GAA AAA GTA GAA Ser Ala Ile His Gln Lys Leu Glu Ala Asp Gly Thr Glu Lys Val Glu	904
	220 225 230 235	
20	GGA TCC ATG ACA CAG AAA TTG GAG AAT GTC CTC AAC AGA GCA AGT AAT Gly Ser Met Thr Gln Lys Leu Glu Asn Val Leu Asn Arg Ala Ser Asn	952
	240 245 250	
25	ACT GCT GAC ACA TTG TTT CAA GAA GTG TTA GGT CGG AAA GAC AAG GCA Thr Ala Asp Thr Leu Phe Gln Glu Val Leu Gly Arg Lys Asp Lys Ala	1000
	255 260 265	
30	GAT TCC ACT AGG AAT GCA CTC AAC GTG CTT CAG CGG TTT AAA TTT CTC Asp Ser Thr Arg Asn Ala Leu Asn Val Leu Gln Arg Phe Lys Phe Leu	1048
	270 275 280	
35	TTC AAC CTT CCT CTC AAT ATC AAA CGG AAC ATT CAA AAG GGG GAT TAT Phe Asn Leu Pro Leu Asn Ile Lys Arg Asn Ile Gln Lys Gly Asp Tyr	1096
	285 290 295	
	GAT GTG GTT ATT AAT GAT TAT GAA AAA GCC AAA TCG CTC TTT GGG AAA Asp Val Val Ile Asn Asp Tyr Glu Lys Ala Lys Ser Leu Phe Gly Lys	1144
	300 305 310 315	
40	ACG GAG GTG CAA GTT TTC AAG AAA TAT TAT GCT GAA GTA GAG GCA GGA Thr Glu Val Gln Val Phe Lys Lys Tyr Tyr Ala Glu Val Glu Ala Gly	1192
	320 325 330	
45	ATT GAA GAT CTA AGA GAG TTA CTT CTA AAG AAA CTG CTC GAG ACT CCG Ile Glu Asp Leu Arg Glu Leu Leu Leu Lys Lys Leu Leu Glu Thr Pro	1240

		112		
	335	340	345	
	TCA ACC TTA CAT GAC CAG AAG CGT TAC ATA AGG TAT CTT TCT GAC CTC	1288		
	Ser Thr Leu His Asp Gln Lys Arg Tyr Ile Arg Tyr Leu Ser Asp Leu			
5	350	355	360	
	CAT GCT CCT GGT GAT CCT GCT TGG CAG TGT ATC GGA GCT CAG CAC AAG	1336		
	His Ala Pro Gly Asp Pro Ala Trp Gln Cys Ile Gly Ala Gln His Lys			
	365	370	375	
10				
	TGG ACC CTC AAA CTC ATG CAA GAC TGC AAA GAA GGC CAC ATG AAG AGC	1384		
	Trp Thr Leu Lys Leu Met Gln Asp Cys Lys Glu Gly His Met Lys Ser			
	380	385	390	395
15				
	CTG AAA GGT AAC CCA GGC CCA CAC AGC CCC ATG CTT GAC CTA GAT AAT	1432		
	Leu Lys Gly Asn Pro Gly Pro His Ser Pro Met Leu Asp Leu Asp Asn			
	400	405	410	
20				
	GAT GCC CGC CCC TCT GTG TTG GGC CAT CTC AGT CAG ACA GCA TCA CTG	1480		
	Asp Ala Arg Pro Ser Val Leu Gly His Leu Ser Gln Thr Ala Ser Leu			
	415	420	425	
	AAG AGA GGC AGC AGC TTC CAG TCT GGT CGA GAT GAC ACA TGG AGG TAC	1528		
25	Lys Arg Gly Ser Ser Phe Gln Ser Gly Arg Asp Asp Thr Trp Arg Tyr			
	430	435	440	
	AAG ACG CCC CAC AGA GTG GCA TTT GTT GAA AAA TTA ACG AAG CTT GTG	1576		
	Lys Thr Pro His Arg Val Ala Phe Val Glu Lys Leu Thr Lys Leu Val			
30	445	450	455	
	TTA AGT CAG CTG CCT AAC TTC TGG AAA CTC TGG ATT TCA TAT GTT AAT	1624		
	Leu Ser Gln Leu Pro Asn Phe Trp Lys Leu Trp Ile Ser Tyr Val Asn			
	460	465	470	475
35				
	GGA AGT CTG TTC AGT GAG ACA GCT GAG AAG TCA GGC CAG ATT GAA AGA	1672		
	Gly Ser Leu Phe Ser Glu Thr Ala Glu Lys Ser Gly Gln Ile Glu Arg			
	480	485	490	
40				
	TCA AAG AAT GTG AGG CAA AGA CAA AAT GAT TTC AAG AAA ATG ATT CAA	1720		
	Ser Lys Asn Val Arg Gln Arg Gln Asn Asp Phe Lys Lys Met Ile Gln			
	495	500	505	
	GAA GTA ATG CAC TCC CTG GTG AAG TTG ATC CGG GGA GCC CTG CTC CCA	1768		
45	Glu Val Met His Ser Leu Val Lys Leu Ile Arg Gly Ala Leu Leu Pro			

113

	510	515	520	
	TTC AGC CTC CGA GAA GGG GAT GGG AGA CAG, TAT GGC GGC TGG GAG GTG			1816
	Phe Ser Leu Arg Glu Gly Asp Gly Arg Gln Tyr Gly Gly Trp Glu Val			
5	525	530	535	
	CAG GCA GAG CTC TCG GGG CAG TGG CTC GCA CAC GTT ATC CAG ACC ATA			1864
	Gln Ala Glu Leu Ser Gly Gln Trp Leu Ala His Val Ile Gln Thr Ile			
	540	545	550	555
10				
	AGG CTG ACA TAT GAA TCA CTG ACC GCT CTG GAG ATT CCC AAT GAC ATG			1912
	Arg Leu Thr Tyr Glu Ser Leu Thr Ala Leu Glu Ile Pro Asn Asp Met			
	560	565	570	
15				
	TTA CAG ATT ATC CAG GAC CTG ATT TTG GAT CTC CGT ATA CAT TGC ATA			1960
	Leu Gln Ile Ile Gln Asp Leu Ile Leu Asp Leu Arg Ile His Cys Ile			
	575	580	585	
20				
	ATG GTG ACG TTG CAA CAT ACA GCT GAA GAG ATA AAG AGA TTA GCT GAA			2008
	Met Val Thr Leu Gln His Thr Ala Glu Glu Ile Lys Arg Leu Ala Glu			
	590	595	600	
	AAA GAA GAT TGG ATT GTT GAT AAC GAA GGA CTG ACT TCC CTA CCA TGT			2056
25	Lys Glu Asp Trp Ile Val Asp Asn Glu Gly Leu Thr Ser Leu Pro Cys			
	605	610	615	
	CAG TTT GAA CAG AGC ATT GTT CAC TCC CTG CAG TCG TTG AAG GGT GTT			2104
	Gln Phe Glu Gln Ser Ile Val His Ser Leu Gln Ser Leu Lys Gly Val			
30	620	625	630	635
	GTG GAT TGC AAA CCT GGA GAA GCC AGT GTC TTT CAG CAG CCT AAA ACC			2152
	Val Asp Cys Lys Pro Gly Glu Ala Ser Val Phe Gln Gln Pro Lys Thr			
	640	645	650	
35				
	CAG GAG GAG GTT TGC CAG TTG TGC ATC AGT ATC ATG CAG GTT TTT ATA			2200
	Gln Glu Glu Val Cys Gln Leu Cys Ile Ser Ile Met Gln Val Phe Ile			
	655	660	665	
40				
	TAC TGC CTG GAA CAA CTG AGC ACC AAG CCT GAT GCG GAC ATA GAT ACT			2248
	Tyr Cys Leu Glu Gln Leu Ser Thr Lys Pro Asp Ala Asp Ile Asp Thr			
	670	675	680	
	ACT CAT CTT TCT GTT GAT GTT TCT TCT CCT GAT TTG TTT GGA AGC ATC			2296
45	Thr His Leu Ser Val Asp Val Ser Ser Pro Asp Leu Phe Gly Ser Ile			

114

	685	690	695	
	CAT GAA GAC TTC GGT TTG ACT TCA GAG CAG AGG CTT TTG ATA GTC CTG			2344
	His Glu Asp Phe Gly Leu Thr Ser Glu Gln Arg Leu Leu Ile Val Leu			
5	700	705	710	715
	AGT AAT TGC TGC TAT TTA GAA CGT CAC ACC TTC CTA AAC ATA GCA GAA			2392
	Ser Asn Cys Cys Tyr Leu Glu Arg His Thr Phe Leu Asn Ile Ala Glu			
10	720	725	730	
	CAT TTT GAA AAG CAC AAC TTC CAG GGA ATA GAA AAA ATA ACA CAG GTT			2440
	His Phe Glu Lys His Asn Phe Gln Gly Ile Glu Lys Ile Thr Gln Val			
	735	740	745	
15				
	AGT ATG GCA TCA CTG AAA GAA CTG GAT CAA CGA CTC TTT GAA AAT TAC			2488
	Ser Met Ala Ser Leu Lys Glu Leu Asp Gln Arg Leu Phe Glu Asn Tyr			
	750	755	760	
20				
	ATT GAG CTA AAA GCA GAT CCT ATT GTT GGA TCC TTG GAA ACT GGG ATA			2536
	Ile Glu Leu Lys Ala Asp Pro Ile Val Gly Ser Leu Glu Thr Gly Ile			
	765	770	775	
25				
	TAT GCA GGC TAT TTT GAT TGG AAA GAC TGT CTG CCT CCA GCA GGT GTC			2584
	Tyr Ala Gly Tyr Phe Asp Trp Lys Asp Cys Leu Pro Pro Ala Gly Val			
	780	785	790	795
30				
	AGA AAC TAT TTA AAA GAA GCA CTG GTG AAT ATA ATT GCT GTG CAT GCA			2632
	Arg Asn Tyr Leu Lys Glu Ala Leu Val Asn Ile Ile Ala Val His Ala			
	800	805	810	
35				
	GAG GTG TTC ACT ATT TCC AAA GAA CTG GTG CCC CGG GTC CTG GCC AGA			2680
	Glu Val Phe Thr Ile Ser Lys Glu Leu Val Pro Arg Val Leu Ala Arg			
	815	820	825	
40				
	GTG ATA GAA GCA GTC TCT GAA GAG CTG AGC CGG CTG ATG CAG TGT GTC			2728
	Val Ile Glu Ala Val Ser Glu Glu Leu Ser Arg Leu Met Gln Cys Val			
	830	835	840	
45				
	TCA TCC TTC AGC AGA AAT GGA GCG CTG CAG GCC AGA CTT GAA ATC TGT			2776
	Ser Ser Phe Ser Arg Asn Gly Ala Leu Gln Ala Arg Leu Glu Ile Cys			
	845	850	855	
	GCT CTG AGG GAC ACT GTA GCC ATT TAC CTG ACC CCG GAG AGC AGG TCT			2824
	Ala Leu Arg Asp Thr Val Ala Ile Tyr Leu Thr Pro Glu Ser Arg Ser			

115

	860	865	870	875	
	AGT TTT AAA CAG GCC CTG GAG GCC CTG CCT CAG CTT GCA AGT GGT GCA				2872
5	Ser Phe Lys Gln Ala Leu Glu Ala Leu Pro Gln Leu Ala Ser Gly Ala	880	885	890	
	GAC AAA AAG CTA CTG GAA GAG CTG CTC AAC AAG TTC AAG AGC AGT ATG				2920
10	Asp Lys Lys Leu Leu Glu Glu Leu Leu Asn Lys Phe Lys Ser Ser Met	895	900	905	
	CAC TTG CAG CTC ACC TGC TTC CAA GCT GCG TCT CCA ACA GTG ATG AAG				2968
	His Leu Gln Leu Thr Cys Phe Gln Ala Ala Ser Pro Thr Val Met Lys	910	915	920	
15	ACA TAGACACAGA CAGAGATCCC AGTCCAAGCA CTGCATTTAC CTTCTCTCTT TTCCCC				3027
	Thr				
	AGCATGGTGG TGCAGAAAGA TGTGATGGG TATTTAATTT CAGCCGATTT AGACTACCAT				3087
20	ACAATTTTGT TATTACTAGC TTTGGTCCTG TAGCATACTC TTCAAATGT TTTTCTTAG				3147
	CTCTATGTTT CTGGGTGGA GTTTTGAGAG CTTCTCCAGC ATGGCCCAAA GCGGCCNATG				3207
	GTGTGTGTGA ATGTGCTNCG CAGTGGCGCT CTACCTGGCC TCAGTGTGG CAGTTTCCAG				3267
	CCCAGGGGAG GATATGAAGT CACAGCTTTT GGAGCTGTCA TAACTTTAA ACCCACCATT				3327
	TCAAACAAGT GGTTCAGGT GCGCTCTCTC TCTCTCTCTC TCTCTCTCTC TCTCTCTCTC				3387
25	TGTAAAGAAA AATATAAAAA TGTTAAATC TGTGATGCAG AAGTCCCTGG ATGAATCTG				3447
	AATTTCATAG TCTGGCTTTC ATTTTCCCA GTGTTCAACA CGGAAGCTTC TGGTTCTCAC				3507
	CCTTAAGTTG ATGCATCGTT GTGGGGAGGG TTTACAAGGA CCTGAATGTG GAGGCCCTGT				3567
	CACACCCGCA AGCCCTGTGG CAGGCTTGTG GTCTTCAATG CAACCTAAAG ATTAAGCCAG				3627
	GCTCCTCCCC TCCCGCAGGT TACTCCCCTC TGCTACTTTC CACACTTCGA AATGTCTTGC				3687
30	TCCATAGCTG GAAAGATGGT GCCCTGGCAC AGGCTGTGTT TTGTAAAGAT TAGAGGTGCC				3747
	TCGGTGACCA GGCATTGAGC AGTGACTCAC TGCTCACTGG CTGGCACAGT AGAATCTGGT				3807
	CACTGCGACC TGCAAGCAGC TTTTACCTCA CTGCATCTGA CCCTAGCTTC AGCTCTGGTT				3867
	CACTCCAGAC AGATCTGTGG CGCCTTTCTT TCAGTTGACA AAATTCCTGT CACTTTTAST				3927
	TCAGGTAATT CCAGAGCACA GCATATGCAC ACTATGGATT ATTTGAGGTT AATAACAGTG				3987
35	TGTTTATTTA TCACAGTAAC TGGAAGTACA AATTTAATCC TTTGTTTGCA AAGTATGTAT				4047
	GTACAGCATT GTCTCCTCAG TGCAGTGTA TTACATCTGG CATAGCCAAA CTGAATTCCA				4107
	AGGCAGAAGT ACACAACTAT TGGTTGTTTA AACACCTAAC AAATACACTA CACAAAGCCA				4167
	GTAGTCTTCC TTCTCTGTAA AACCTCTAAG TGTCTTTATT TCACCATTTT TTATTAGCTC				4227
	TGTAATTGTT TATATTCCTA GTTTTAAAT CCAAATAAAA TAGTGGTGTT CTCTCTTTAA				4287
40	AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA				4347
	AAAAAAAAAA A				4358

(2) INFORMATION FOR SEQ ID NO:62:



116

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 924 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

5

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

```

10  Met Ser Arg Ser Arg Gln Pro Pro Leu Val Thr Gly Ile Ser Pro Asn
    1           5           10           15
    Glu Gly Ile Pro Trp Thr Lys Val Thr Ile Arg Gly Glu Asn Leu Gly
        20           25           30
    Thr Gly Pro Thr Asp Leu Ile Gly Leu Thr Ile Cys Gly His Asn Cys
15      35           40           45
    Leu Leu Thr Ala Glu Trp Met Ser Ala Ser Lys Ile Val Cys Arg Val
        50           55           60
    Gly Gln Ala Lys Asn Asp Lys Gly Asp Ile Ile Val Thr Thr Lys Ser
        65           70           75           80
20  Gly Gly Arg Gly Thr Ser Thr Val Ser Phe Lys Leu Leu Lys Pro Glu
        85           90           95
    Lys Ile Gly Ile Leu Asp Gln Ser Ala Val Trp Val Asp Glu Met Asn
        100          105          110
    Tyr Tyr Asp Met Arg Thr Asp Arg Asn Lys Gly Ile Pro Pro Leu Ser
25      115          120          125
    Leu Arg Pro Ala Asn Pro Leu Gly Ile Glu Ile Glu Lys Cys Lys Leu
        130          135          140
    Pro Gln Lys Asn Leu Glu Val Leu Phe His Gly Met Ser Ala Asp Phe
        145          150          155          160
30  Thr Ser Glu Asn Phe Ser Ala Ala Trp Tyr Leu Ile Glu Asn His Ser
        165          170          175
    Asn Thr Ser Phe Glu Gln Leu Lys Met Ala Val Thr Asn Leu Lys Arg
        180          185          190
    Gln Ala Asn Lys Lys Ser Glu Gly Ser Leu Ala Tyr Val Lys Gly Gly
35      195          200          205
    Leu Ser Thr Phe Phe Glu Ala Gln Asp Ala Leu Ser Ala Ile His Gln
        210          215          220
    Lys Leu Glu Ala Asp Gly Thr Glu Lys Val Glu Gly Ser Met Thr Gln
        225          230          235          240
40  Lys Leu Glu Asn Val Leu Asn Arg Ala Ser Asn Thr Ala Asp Thr Leu
        245          250          255
    Phe Gln Glu Val Leu Gly Arg Lys Asp Lys Ala Asp Ser Thr Arg Asn
        260          265          270
    Ala Leu Asn Val Leu Gln Arg Phe Lys Phe Leu Phe Asn Leu Pro Leu
45      275          280          285

```

117

	Asn	Ile	Lys	Arg	Asn	Ile	Gln	Lys	Gly	Asp	Tyr	Asp	Val	Val	Ile	Asn
	290						295					300				
	Asp	Tyr	Glu	Lys	Ala	Lys	Ser	Leu	Phe	Gly	Lys	Thr	Glu	Val	Gln	Val
	305					310					315				320	
5	Phe	Lys	Lys	Tyr	Tyr	Ala	Glu	Val	Glu	Ala	Gly	Ile	Glu	Asp	Leu	Arg
						325					330				335	
	Glu	Leu	Leu	Leu	Lys	Lys	Leu	Leu	Glu	Thr	Pro	Ser	Thr	Leu	His	Asp
						340				345				350		
	Gln	Lys	Arg	Tyr	Ile	Arg	Tyr	Leu	Ser	Asp	Leu	His	Ala	Pro	Gly	Asp
10			355					360					365			
	Pro	Ala	Trp	Gln	Cys	Ile	Gly	Ala	Gln	His	Lys	Trp	Thr	Leu	Lys	Leu
	370						375					380				
	Met	Gln	Asp	Cys	Lys	Glu	Gly	His	Met	Lys	Ser	Leu	Lys	Gly	Asn	Pro
	385					390					395				400	
15	Gly	Pro	His	Ser	Pro	Met	Leu	Asp	Leu	Asp	Asn	Asp	Ala	Arg	Pro	Ser
						405					410				415	
	Val	Leu	Gly	His	Leu	Ser	Gln	Thr	Ala	Ser	Leu	Lys	Arg	Gly	Ser	Ser
						420				425				430		
	Phe	Gln	Ser	Gly	Arg	Asp	Asp	Thr	Trp	Arg	Tyr	Lys	Thr	Pro	His	Arg
20			435					440					445			
	Val	Ala	Phe	Val	Glu	Lys	Leu	Thr	Lys	Leu	Val	Leu	Ser	Gln	Leu	Pro
	450						455					460				
	Asn	Phe	Trp	Lys	Leu	Trp	Ile	Ser	Tyr	Val	Asn	Gly	Ser	Leu	Phe	Ser
	465					470					475				480	
25	Glu	Thr	Ala	Glu	Lys	Ser	Gly	Gln	Ile	Glu	Arg	Ser	Lys	Asn	Val	Arg
						485					490			495		
	Gln	Arg	Gln	Asn	Asp	Phe	Lys	Lys	Met	Ile	Gln	Glu	Val	Met	His	Ser
						500				505				510		
30	Leu	Val	Lys	Leu	Ile	Arg	Gly	Ala	Leu	Leu	Pro	Phe	Ser	Leu	Arg	Glu
			515					520					525			
	Gly	Asp	Gly	Arg	Gln	Tyr	Gly	Gly	Trp	Glu	Val	Gln	Ala	Glu	Leu	Ser
	530						535					540				
	Gly	Gln	Trp	Leu	Ala	His	Val	Ile	Gln	Thr	Ile	Arg	Leu	Thr	Tyr	Glu
35	545					550					555				560	
	Ser	Leu	Thr	Ala	Leu	Glu	Ile	Pro	Asn	Asp	Met	Leu	Gln	Ile	Ile	Gln
						565					570			575		
	Asp	Leu	Ile	Leu	Asp	Leu	Arg	Ile	His	Cys	Ile	Met	Val	Thr	Leu	Gln
						580				585				590		
40	His	Thr	Ala	Glu	Glu	Ile	Lys	Arg	Leu	Ala	Glu	Lys	Glu	Asp	Trp	Ile
			595					600					605			
	Val	Asp	Asn	Glu	Gly	Leu	Thr	Ser	Leu	Pro	Cys	Gln	Phe	Glu	Gln	Ser
			610					615					620			
	Ile	Val	His	Ser	Leu	Gln	Ser	Leu	Lys	Gly	Val	Val	Asp	Cys	Lys	Pro
45	625					630					635			640		

118

	Gly	Glu	Ala	Ser	Val	Phe	Gln	Gln	Pro	Lys	Thr	Gln	Glu	Glu	Val	Cys
					645					650					655	
	Gln	Leu	Cys	Ile	Ser	Ile	Met	Gln	Val	Phe	Ile	Tyr	Cys	Leu	Glu	Gln
				660					665					670		
5	Leu	Ser	Thr	Lys	Pro	Asp	Ala	Asp	Ile	Asp	Thr	Thr	His	Leu	Ser	Val
				675					680					685		
	Asp	Val	Ser	Ser	Pro	Asp	Leu	Phe	Gly	Ser	Ile	His	Glu	Asp	Phe	Gly
				690				695					700			
	Leu	Thr	Ser	Glu	Gln	Arg	Leu	Leu	Ile	Val	Leu	Ser	Asn	Cys	Cys	Tyr
10	705					710					715					720
	Leu	Glu	Arg	His	Thr	Phe	Leu	Asn	Ile	Ala	Glu	His	Phe	Glu	Lys	His
					725					730					735	
	Asn	Phe	Gln	Gly	Ile	Glu	Lys	Ile	Thr	Gln	Val	Ser	Met	Ala	Ser	Leu
				740					745					750		
15	Lys	Glu	Leu	Asp	Gln	Arg	Leu	Phe	Glu	Asn	Tyr	Ile	Glu	Leu	Lys	Ala
				755					760					765		
	Asp	Pro	Ile	Val	Gly	Ser	Leu	Glu	Thr	Gly	Ile	Tyr	Ala	Gly	Tyr	Phe
				770					775				780			
	Asp	Trp	Lys	Asp	Cys	Leu	Pro	Pro	Ala	Gly	Val	Arg	Asn	Tyr	Leu	Lys
20	785					790					795					800
	Glu	Ala	Leu	Val	Asn	Ile	Ile	Ala	Val	His	Ala	Glu	Val	Phe	Thr	Ile
					805						810				815	
	Ser	Lys	Glu	Leu	Val	Pro	Arg	Val	Leu	Ala	Arg	Val	Ile	Glu	Ala	Val
				820					825					830		
25																
	Ser	Glu	Glu	Leu	Ser	Arg	Leu	Met	Gln	Cys	Val	Ser	Ser	Phe	Ser	Arg
				835				840						845		
	Asn	Gly	Ala	Leu	Gln	Ala	Arg	Leu	Glu	Ile	Cys	Ala	Leu	Arg	Asp	Thr
				850				855						860		
30	Val	Ala	Ile	Tyr	Leu	Thr	Pro	Glu	Ser	Arg	Ser	Ser	Phe	Lys	Gln	Ala
	865					870						875				880
	Leu	Glu	Ala	Leu	Pro	Gln	Leu	Ala	Ser	Gly	Ala	Asp	Lys	Lys	Leu	Leu
					885					890					895	
	Glu	Glu	Leu	Leu	Asn	Lys	Phe	Lys	Ser	Ser	Met	His	Leu	Gln	Leu	Thr
35				900					905					910		
	Cys	Phe	Gln	Ala	Ala	Ser	Pro	Thr	Val	Met	Lys	Thr				
				915					920							

## IT IS CLAIMED:

1. A substantially purified polypeptide selected from the group consisting of
  - 5 a. SA-17S p71, having a molecular weight of about 71 kD and having the sequence of SEQ ID NO:54;
  - b. SA-17S p79, having a molecular weight of about 79 kD and having the sequence of SEQ ID NO:56;
  - c. SA-17S p84, having a molecular weight of about 84 kD and having the sequence of SEQ ID NO:58;
  - 10 d. SA-17S p96, having a molecular weight of about 96 kD and having the sequence of SEQ ID NO:60;
  - e. SA-17S p102, having a molecular weight of about 102 kD and having the sequence of SEQ ID NO:62;
  - 15 f. SA-17S p106, having a molecular weight of about 106 kD and characterized by sequences selected from the group consisting of SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, and SEQ ID NO:52; and
  - g. a polypeptide having (i) at least 60% sequence identity with a selected one of the SA-17S polypeptides (a)-(f), and (ii) the property of the selected polypeptide to form an SC complex-binding SA-17S complex when combined stoichiometrically with other polypeptide components of the SA-17S complex.
2. The polypeptide of claim 1, which is selected from the group consisting of (a)-(e) and (h),
  - h. a polypeptide having (i) at least 60% sequence identity with a selected one of the SA-17S polypeptides (a)-(e), and (ii) the property of the selected peptide to form an SC complex-binding SA-17S complex when combined stoichiometrically with other polypeptide components of the SA-17S complex.
3. The polypeptide of claim 2, which is polypeptide (a) or a polypeptide having (i) at least 60% sequence identity with polypeptide (a) and the property of the polypeptide (a) to form an SC complex-binding SA-17S complex when

combined stoichiometrically with other polypeptide components of the SA-17S complex.

4. The polypeptide of claim 2, which is polypeptide (b) or a polypeptide having (i) at least 60% sequence identity with polypeptide (b) and the property of the polypeptide (b) to form an SC complex-binding SA-17S complex when combined stoichiometrically with other polypeptide components of the SA-17S complex.

5. The polypeptide of claim 2, which is polypeptide (c) or a polypeptide having (i) at least 60% sequence identity with polypeptide (c) and the property of the polypeptide (c) to form an SC complex-binding SA-17S complex when combined stoichiometrically with other polypeptide components of the SA-17S complex.

6. The polypeptide of claim 2, which is polypeptide (d) or a polypeptide having (i) at least 60% sequence identity with polypeptide (d) and the property of the polypeptide (d) to form an SC complex-binding SA-17S complex when combined stoichiometrically with other polypeptide components of the SA-17S complex.

7. The polypeptide of claim 2, which is polypeptide (e) or a polypeptide having (i) at least 60% sequence identity with polypeptide (e) and the property of the polypeptide (e) to form an SC complex-binding SA-17S complex when combined stoichiometrically with other polypeptide components of the SA-17S complex.

8. A substantially purified polynucleotide encoding one of the polypeptides (a)-(g) in claim 1, said polynucleotide being selected from the group consisting of

a. a polynucleotide having the sequence SEQ ID NO:53, which encodes the SA-17S p71 protein;

b. a polynucleotide having the sequence SEQ ID NO:55, which encodes the SA-17S p79 protein;

- c. a polynucleotide having the sequence SEQ ID NO:57, which encodes the SA-17S p84 protein;
- d. a polynucleotide sequence having the sequence SEQ ID NO:59, which encodes the SA-17S p96 protein;
- 5 e. a polynucleotide having the sequence SEQ ID NO:61, which encodes the SA-17S p102 protein; and
- f. a polynucleotide (i) having at least 60% sequence identity with a selected one of the SA-17S polynucleotides (a)-(e), and (ii) encoding a polypeptide having the property of the polypeptide encoded by the selected polynucleotide to form an
- 10 SC complex-binding SA-17S complex when combined stoichiometrically with other polypeptide components of the SA-17S complex.
9. The polynucleotide of claim 8, which is polynucleotide (a) or a polynucleotide having (i) at least 60% sequence identity with polynucleotide (a) and
- 15 the property of the SA-17S p71 polypeptide encoded by polynucleotide (a) to form an SC complex-binding SA-17S complex when combined stoichiometrically with other polypeptide components of the SA-17S complex.
10. The polynucleotide of claim 8, which is polynucleotide (b) or a polynucleotide having (i) at least 60% sequence identity with polynucleotide (b) and
- 20 the property of the SA-17S p79 polypeptide encoded by polynucleotide (b) to form an SC complex-binding SA-17S complex when combined stoichiometrically with other polypeptide components of the SA-17S complex.
- 25 11. The polynucleotide of claim 8, which is polynucleotide (c) or a polynucleotide having (i) at least 60% sequence identity with polynucleotide (c) and the property of the SA-17S p84 polypeptide encoded by polynucleotide (c) to form an SC complex-binding SA-17S complex when combined stoichiometrically with other polypeptide components of the SA-17S complex.
- 30 12. The polynucleotide of claim 8, which is polynucleotide (d) or a polynucleotide having (i) at least 60% sequence identity with polynucleotide (d) and the property of the SA-17S p96 polypeptide encoded by polynucleotide (d) to form

an SC complex-binding SA-17S complex when combined stoichiometrically with other polypeptide components of the SA-17S complex.

13. The polynucleotide of claim 8, which is polynucleotide (e) or a  
5 polynucleotide having (i) at least 60% sequence identity with polynucleotide (e) and the property of the SA-17S p102 polypeptide encoded by polynucleotide (e) to form an SC complex-binding SA-17S complex when combined stoichiometrically with other polypeptide components of the SA-17S complex.

10 14. A method of identifying a compound capable of modulating secretion of secretory vesicles, comprising  
contacting a secretion associated 17S (SA-17S) complex containing a polypeptide from claim 1 with a syntaxin-containing (SC) complex in the presence and absence of a test compound,  
15 measuring the effect of the test compound on the extent of binding between the SA-17S and the SC complexes, and  
identifying said compound as effective if its measured effect on the extent of binding is above a threshold level.

20 15. The method of claim 14, wherein said SA-17S complex binds to the 7S SC complex or the 20S SC complex.

25 16. The method of claim 14, wherein said secretory vesicles are neurotransmitter-containing synaptic vesicles.

17. The method of claim 16, wherein the test compound is effective to enhance binding of the SC complex with said SA-17S complex.

30 18. The method of claim 16, wherein the test compound is effective to inhibit binding of the SC complex with said SA-17S complex.

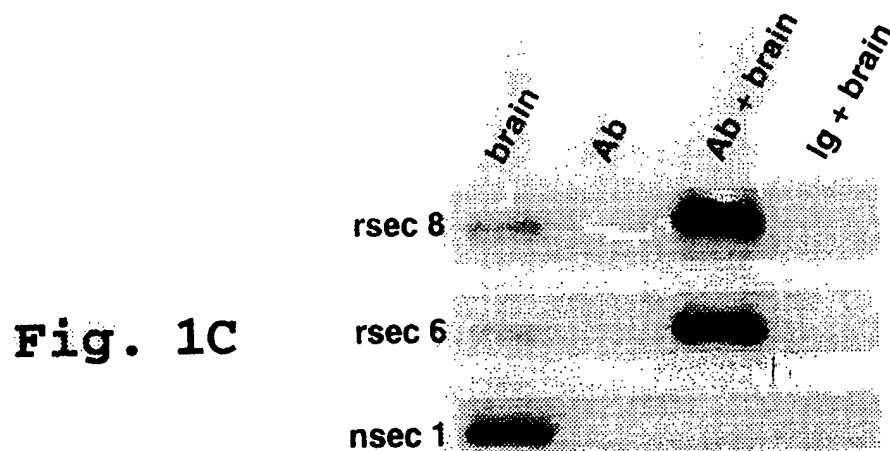
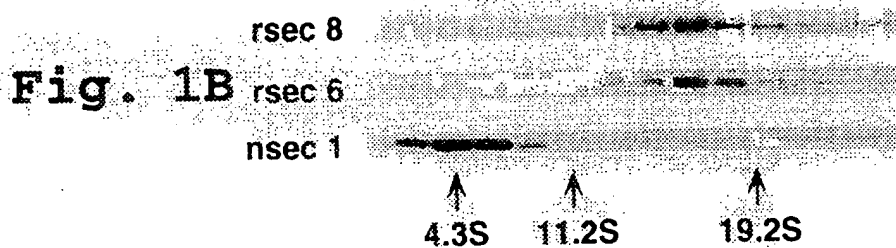
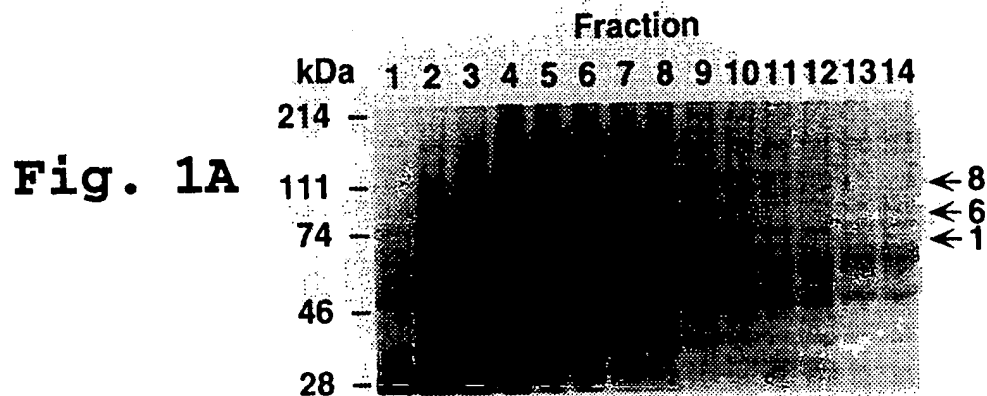
19. A method for treating an affective disorder comprising administering to a subject in need of such treatment an effective amount of a compound identified in claim 17.

5           20. A method for treating a neurodegenerative disease comprising administering to a subject in need of such treatment an effective amount of a compound identified in claim 17.

10           21. A method for treating schizophrenia comprising administering to a subject in need of such treatment an effective amount of a compound identified in claim 18.



1/9



2/9

Fig. 2A

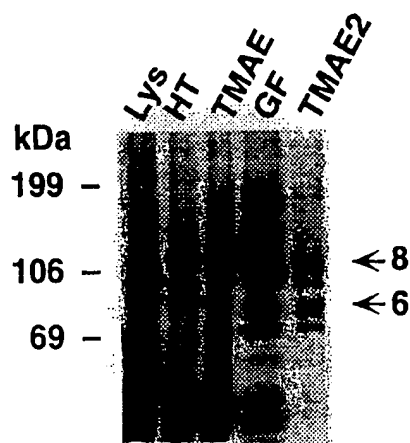
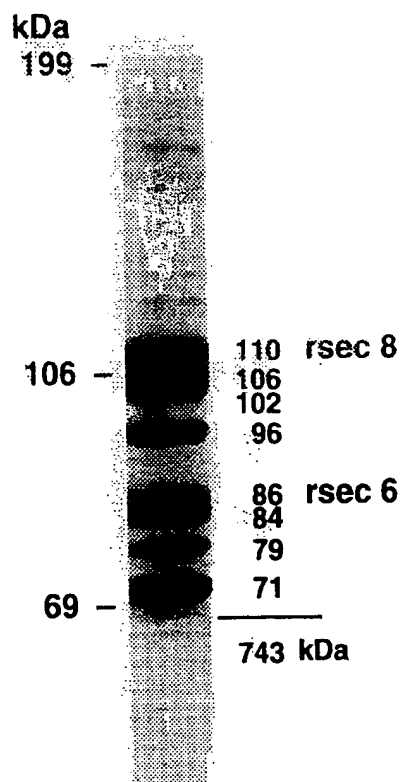


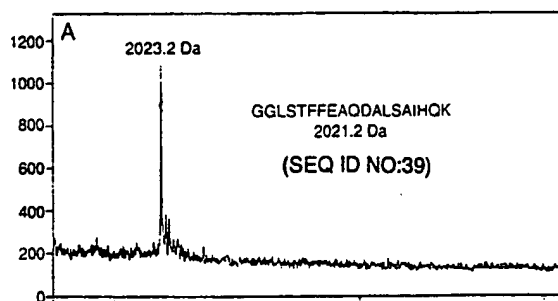
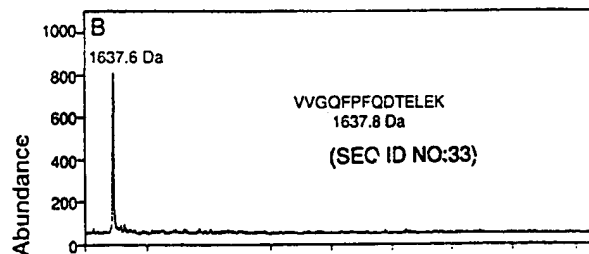
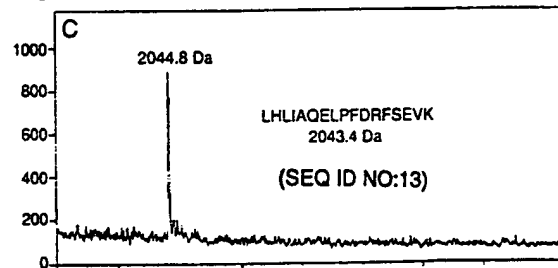
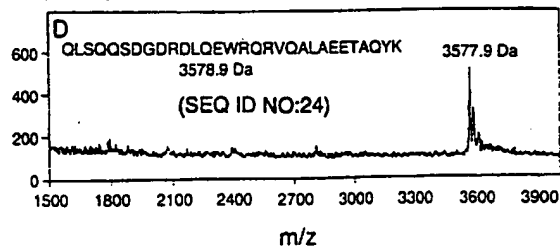
Fig. 2B



Fig. 2C



3/9

**Fig. 3A****Fig. 3B****Fig. 3C****Fig. 3D**

rsec8 1. XAPEGPLIDVXNI  
2. EFAAFFAK  
3. XLGVQRPLLQSTSIXE

**Fig. 4A**

p106 1. XXYGEIAXK  
2. XATVSLPEK  
3. XDYGVIAND  
4. LIKYFFMVASVK  
5. ELPEFNLHFFK  
6. XLQDVDLASXR  
7. XNRXNEPAVNVL  
8. XQLXNIVEPEXIY

**Fig. 4B**

p102 1. YLSGLQAPGXPASQSIGAQ  
2. GGLSTFFEAQDALSAIHQK  
3. ASNTADTXRQER  
4. XRENYIEGGI  
5. ENLGRLFENYI  
6. XDYDVVINDYE  
7. XIPXLSTRPANP

**Fig. 4C**

p96 1. VVGQFPFQDTELEK  
2. XVEIFDN  
3. DFLESIR  
4. XDQDLQLADYDHMT  
5. STNLLLTRLXN

**Fig. 4D**

**Fig. 4E**

rsec6 1. DFRQSINTIEXL  
2. QGPSQASPNYXP  
3. AAIQSQLDGVRTGLSQ

**Fig. 4F**

p84 1. QLSQQSDGDRDLQEWQRQVQALAEETAQYK  
2. XLQLSFNFSEPNRQRP  
3. SIPLALLPAAAAGA  
4. DAVXQNSTQAAETEN  
5. DYRNDEA  
6. ENNPEEDDPS  
7. XLSQQSDXG  
8. AAALRAPPXVTS  
9. XKREPLE

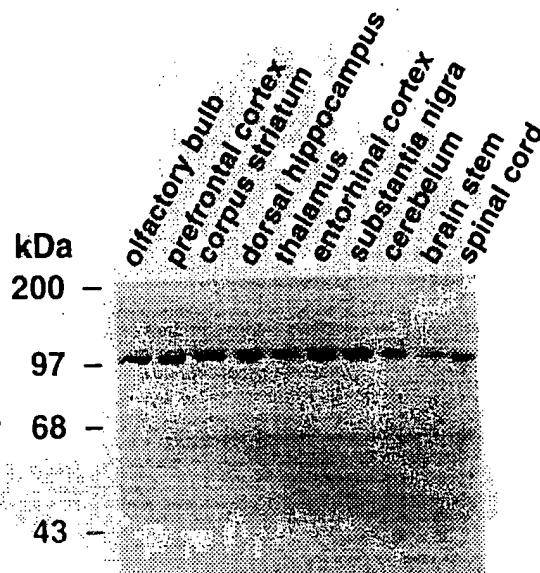
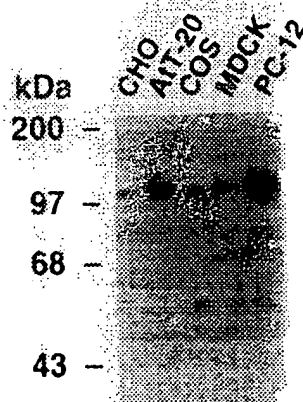
**Fig. 4G**

p79 1. XTDYIAE  
2. ETYGAFLSRSXG  
3. XXPPQGVYPNPASP  
4. XXQEEETLMFIRGN  
5. ALFIRDDXQF  
6. NLPVFQSCS  
7. AVEYFQDKFPD  
8. YRVEQVGDMIDRLFDTS  
9. VYEDPALSAIFLHNNYNY  
10. XXYGAFLHRYSSVPFVYGXH

**Fig. 4H**

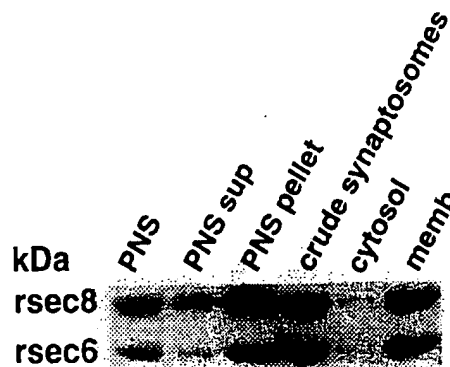
p71 1. XNQVAFQHFQELDEHI  
2. VCHLXDQLEXVN  
3. LHLIAQELPFDRFSEVK

6/9

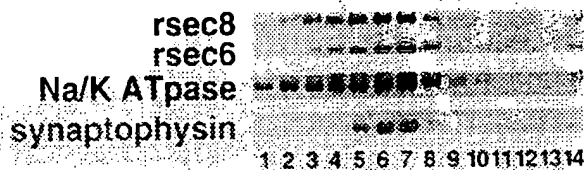
**Fig. 5A****Fig. 5B**

7/9

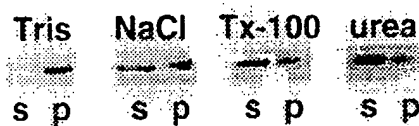
**Fig. 6A**



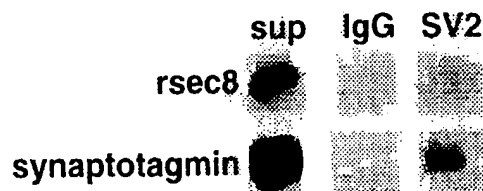
**Fig. 6B**



**Fig. 6C**



**Fig. 6D**



**Fig. 7A**



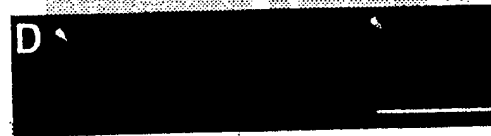
**Fig. 7B**



**Fig. 7C**



**Fig. 7D**





9/9

